PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 18, Fasc. 2 (1998), pp. 219–245

UNE ÉTUDE ALGÉBRIQUE DE L'ADMISSIBILITÉ EN ESTIMATION LINÉAIRE DE LA MOYENNE SUR UN MODÈLE GÉNÉRAL DE GAUSS–MARKOV

PAR

JEAN-JACQUES TÉCHENÉ (PAU, FRANCE)

Abstract. It would appear useful to come back to the question of admissibility in linear estimation on a general Gauss–Markov model. We prove how a functional approach to this problem, based on a very important LaMotte theorem [11], clearly leads to characterization of all admissible linear estimators of mean vector or linear transformation of mean vector. Thus we have managed to modify significantly a Klonecki and Zontek theorem [9] allowing us to find in a different way an essential characterization shown by Baksalary and Markiewicz [4], based on the logic put forward by Rao (cf. [13] and [14]). We also give a variational characterization of admissibility in linear estimation and a geometrical proof of a Baksalary and Mathew theorem [7] relative to equality between the set of best linear unbiased estimators (or Gauss–Markov estimators) and the set of linear admissible estimators of mean vector. We finish by explaining more results on admissibility of linear estimators of vector parameters.

INTRODUCTION

Cet article est une suite de [18] et de [20]. La structure statistique considérée est maintenant un modèle général de Gauss–Markov, autrement dit une structure statistique linéaire induite sur un espace vectoriel réel de dimension finie E par un vecteur aléatoire Y (sommable à l'ordre deux) tel que l'ensemble des couples (μ_θ, V_θ) formés de l'espérance mathématique μ_θ et du tenseur de variance V_θ de Y, pour toutes les probabilités P_θ mises dans le modèle, soit un *produit cartésien* $L \times \{V\}$ donné, où V est un morphisme symétrique et positif du dual E^* dans E (non nécessairement bijectif) et L un sous-espace vectoriel ou affine de E. *L'espace E n'est pas nécessairement un espace R^n* (modèle univarié): *ce peut être aussi un espace d'applications linéaires* (modèle multivarié). Nous montrons comment le langage fonctionnel et de la géométrie affine facilite dans ce cadre, par la mise en oeuvre de la structure algébrique induite par V sur E (structure étudiée dans [19]), une étude de l'admissibilité en estimation linéaire de toute transformation linéaire de la moyenne μ de Y. A l'instar de Klonecki et Zontek [9], notre approche se fonde sur un théorème capital de LaMotte (cf. [11] et [18]) rattachant la théorie de l'admissibilité en estimation linéaire à la résolution de programmes quadratiques sous contraintes linéaires ou affines permettant de dégager efficacement des conditions d'admissibilité.

Nous obtenons, dans ce cadre d'un modèle de Gauss–Markov non nécessairement univarié, une version générale et plus précise d'un théorème caractéristique dû à Klonecki et Zontek [9] que nous relions à un autre théorème caractéristique important établi par Baksalary et Markiewicz [4] par une voie totalement différente se situant sur une lignée importante de travaux ouverte par Rao dans [17]. Nous en déduisons une caractérisation variationnelle de l'admissibilité en estimation linéaire et montrons comment on peut reconstituer algébriquement la famille des estimateurs linéaires admissibles de chaque transformation linéaire de la moyenne μ à partir des estimateurs linéaires admissibles de μ. Nous donnons également une preuve géométrique, mettant encore en oeuvre la structure algébrique induite sur E par le morphisme V, d'un théorème remarquable dû à Baksalary et Mathew [7] dans le cas univarié, indiquant que les ensembles des estimateurs linéaires admissibles et des estimateurs de Gauss–Markov de la moyenne μ coïncident si et seulement si le sous-espace $L \cap \operatorname{Im} V$ est réduit au vecteur nul.

Nous terminons par un approfondissement de la caractérisation algébrique des estimateurs linéaires ou affines admissibles du paramètre b de régression dans un modèle général de Gauss–Markov (éventuellement multivarié). Nous donnons une formulation plus complète d'une propriété capitale découverte par Rao [17] lorsque E est un espace $\boldsymbol{R}^n$ (modèle univarié), indiquant que tout estimateur linéaire admissible de b est l'image d'un estimateur des moindres-carrés généralisés de b par une application linéaire: nous spécifions le caractère statistique de ce morphisme. Lorsque le modèle est placé sous contraintes structurelles affines, nous établissons plusieurs propriétés caractéristiques de la famille des estimateurs linéaires ou affines admissibles de b qui apparaissent comme des versions descriptives plus générales et plus précises de résultats dus à Baksalary et Markiewicz [2] dans le cas univarié.

1. LES PRINCIPES FONDAMENTAUX DE L'ESTIMATION LINÉAIRE

1.1. Notations et généralités. Nous utilisons les mêmes notations que dans [20]. Pour tout couple (E, F) d'espaces vectoriels réels, nous désignerons par $\mathscr{L}(E, F)$ [resp. $\mathscr{L}(E)$] l'espace des applications linéaires de E dans F [resp. E]. Pour chaque application linéaire T de E dans F et toute partie A de E, le noyau, l'image, la transposée, le rang et la restriction de T au sous-ensemble A (autrement dit, l'application de A dans F qui à tout x de A associe Tx) seront notés $\operatorname{Ker} T$, $\operatorname{Im} T$, ${}^tT$, $\operatorname{Rg}(T)$ et T_A respectivement. Le symbole o désigne la loi de composition des applications.

On appelle *g-inverse* d'une application linéaire T de E dans F toute application linéaire T^- de F dans E telle que $T \circ T^- \circ T = T$. Il équivaut à dire que $T \circ T^-$ est un projecteur de F sur $\operatorname{Im} T$, ou encore que $I_E - T^- \circ T$ est un projecteur de E sur $\operatorname{Ker} T$, I_E désignant l'application identique de E. Une g-inverse T^- de T est dite *réflexive* si T est une inverse généralisée de T^-, ou ce qui est équivalent, si T^- et $T^- \circ T$ ont même image dans E. Lorsque E et F sont euclidiens, le pseudo-inverse de Moore–Penrose est l'unique g-inverse réflexive $T^\dagger$ de T telle que $T \circ T^\dagger$ et $I_E - T^\dagger \circ T$ soient respectivement le projecteur orthogonal de F sur $\operatorname{Im} T$ et de E sur $\operatorname{Ker} T$.

Pour tout espace vectoriel réel de dimension *finie* E, de dual E^* (espace des formes linéaires sur E), nous noterons $\langle \cdot, \cdot \rangle$ la dualité canonique sur $E^* \times E$ et identifierons l'espace des formes bilinéaires sur $E^* \times E^*$ à l'espace $\mathscr{L}(E^*, E)$ dont les éléments seront appelés *tenseurs sur* E^* (cf. § 2.3 de [20]). L'espace des tenseurs symétriques [resp. positifs] sur E^* sera noté $\mathscr{L}_s(E^*, E)$ [resp. $\mathscr{L}_s^+(E^*, E)$]. Pour tout autre espace vectoriel réel F et chaque couple (Z, W) de sous-espaces vectoriels de E et F respectivement, le produit tensoriel $Z \otimes W$ désigne le sous-espace vectoriel de $\mathscr{L}(E, F)$ engendré par l'ensemble $\{u^* \otimes y \colon (u^*, y) \in Z \times W\}$, où $u^* \otimes y$ est l'application linéaire élémentaire (de rang 1) de E dans F définie par (cf. § 2.2 de [20]):

$$u^* \otimes y \colon x \to \langle u^*, x \rangle y.$$

Considérons deux espaces vectoriels réels E et F de dimension finie, munis de leurs topologies canoniques et de leurs tribus de Borel $\mathscr{B}_E$ et $\mathscr{B}_F$. Soit $(E, \mathscr{B}_E, \mathscr{P})$ une structure statistique sur E induite par un vecteur aléatoire Y sommable à l'ordre deux. Nous supposons (sans perte de généralité) que *Y est la statistique identique de E* et que la famille $\mathscr{P}$ de probabilités est décrite à l'aide d'un paramètre θ variant dans un ensemble connu Θ, ce que l'on écrira $\mathscr{P} = \{P_\theta \colon \theta \in \Theta\}$.

Pour tout couple (S, T) de statistiques à valeurs dans F, nous noterons $S \otimes T$ la statistique à valeurs dans l'espace $\mathscr{L}(F^*, F) = F \otimes F$ (cf. proposition 2.1 de [20]) définie par

$$(\forall x \in E) \ \ (S \otimes T) \colon v^* \to \langle Sx, v^* \rangle \, Tx = (Sx \otimes Tx) \, v^*.$$

La valeur de l'espérance mathématique (ou moyenne) μ de Y pour chaque θ de Θ est définie comme l'unique vecteur $\mu_\theta := \boldsymbol{E}_\theta(Y)$ de E tel que

$$(\forall u^* \in E^*) \ \ \langle u^*, \mu_\theta \rangle = \boldsymbol{E}_\theta(\langle u^*, Y \rangle).$$

Le tenseur de variance

$$V_\theta(Y) = \boldsymbol{E}_\theta[(Y - \mu_\theta) \otimes (Y - \mu_\theta)]$$

de Y est identifié à l'unique application linéaire symétrique (et positive) de E^* dans E, notée V_θ, vérifiant:

$$(\forall (u^*, v^*) \in E^* \times E^*) \ \ \langle v^*, V_\theta u^* \rangle = \operatorname{Cov}_\theta(\langle u^*, Y \rangle, \langle v^*, Y \rangle).$$

Soit U un estimateur sommable à l'ordre deux d'une application g de Θ dans $(F, \mathscr{B}_F)$. Le tenseur $\varrho(\theta, U)$ de l'erreur quadratique de U (erreur commise en estimant g par U) est défini par

$$\varrho(\theta, U) := E_\theta\{[U - g(\theta)] \otimes [U - g(\theta)]\}.$$

Pour comparer deux estimateurs U_1 et U_2 de g, admettant chacun un tenseur de variance, on utilise le préordre partiel $\preccurlyeq_1$, déduit de l'ordre de Loewner $\preccurlyeq_L$ sur l'espace $\mathscr{L}(F^*, F)$ (cf. [20], § 2.3), défini par

$$U_1 \preccurlyeq_1 U_2 \Leftrightarrow (\forall \theta \in \Theta)\ \varrho(\theta, U_1) \preccurlyeq_L \varrho(\theta, U_2),$$

et, si F est mis en dualité séparante avec lui-même (cf. § 1.1 de [20]), le préordre $\preccurlyeq_2$, dérivé du préordre de Schur sur $\mathscr{L}(F)$ (cf. [20], § 2.3), défini par

$$U_1 \preccurlyeq_2 U_2 \Leftrightarrow (\forall \theta \in \Theta)\ \mathrm{tr}[\varrho(\theta, U_1)] \leqslant \mathrm{tr}[\varrho(\theta, U_2)].$$

Etant donnée une famille $\mathscr{G}(g)$ d'estimateurs de g admettant un tenseur de variance, un élément de $\mathscr{G}(g)$ est dit *optimal dans* $\mathscr{G}(g)$ s'il est minimal relativement au préordre $\preccurlyeq_1$, et *admissible dans* $\mathscr{G}(g)$ s'il est minimal relativement au préodre $\preccurlyeq_2$ (cf. [20], § 3.2).

1.2. Les concepts de structure et d'estimateur linéaires. Nous dirons que la structure statistique $(E, \mathscr{B}_E, \mathscr{P})$ est *linéaire* si la moyenne et le tenseur de variance de Y sont *caractéristiques de la famille* $\mathscr{P}$, au sens où la donnée de la famille $\mathscr{P}$ implique la connaissance de l'ensemble $\{(\mu_\theta, V_\theta) \colon \theta \in \Theta\}$, bien que parfois seules les projections canoniques $\{\mu_\theta \colon \theta \in \Theta\}$ et $\{V_\theta \colon \theta \in \Theta\}$ soient à connaître, comme c'est le cas en estimation linéaire sans biais de variance uniformément minimum.

Toutes les statistiques mises en jeu dans toute la suite de cette section sont supposées définies sur une structure linéaire $(E, \mathscr{B}_E, \mathscr{P})$, où $\mathscr{P} = \{P_\theta \colon \theta \in \Theta\}$. Nous conviendrons des définitions suivantes:

(1) Une application affine B de E dans F est la composée $t_u \circ A$ d'une translation t_u de F (de vecteur u) par une application linéaire A de E dans F, dite *partie linéaire de* B.

(2) Pour toute application affine B de E dans F, nous nommerons *estimateur linéaire ou affine de* $B\mu$ (ou tout simplement, *de* B) chaque estimateur linéaire ou affine U de l'application $g \colon \theta \to B\mu_\theta$ de Θ dans F, μ désignant la moyenne de Y.

Si B et U sont linéaires, le tenseur d'erreur quadratique $\varrho(\theta, U)$ de U est donné par

$$\varrho(\theta, U) = U \circ V_\theta \circ {}^tU + (U - B) \circ (\mu_\theta \otimes \mu_\theta) \circ {}^t(U - B),$$

et, si U est un estimateur linéaire sans biais de B (autrement dit, vérifiant $E_\theta(U) = B\mu_\theta$ pour chaque θ de Θ), par

$$\varrho(\theta, T) = U \circ V_\theta \circ {}^tU.$$

Il apparaît donc clairement que l'espace paramétrique "naturel" associé à l'estimation linéaire (ou affine) sur une structure linéaire est le sous-ensemble

$$\Sigma = \{(V_\theta, \mu_\theta \otimes \mu_\theta): \theta \in \Theta\}$$

de l'espace-produit $\mathscr{L}_s(E^*, E) \times \mathscr{L}_s(E^*, E)$. Relevons que Σ n'est pas nécessairement un produit cartésien et que la restriction des relations de comparaison de deux estimateurs linéaires de B à la classe des estimateurs linéaires *sans biais* de B *n'exige que la connaissance des deux projections canoniques de* Σ.

Selon l'usage, nous dirons que la structure linéaire $(E, \mathscr{B}_E, \mathscr{P})$ est un *modèle général de Gauss–Markov* (éventuellement multivarié) si $\{(\mu_\theta, V_\theta): \theta \in \Theta\}$ est un *produit cartésien* de la forme $L \times \{V\}$, où V est un élément non nul de $\mathscr{L}_{es}^+(E^*, E)$ et L un sous-espace vectoriel ou affine de E donnés, ce qui équivaut à supposer que

$$\Sigma = \{V\} \times \{u \otimes u: \ u \in L\}.$$

Une telle hypothèse est importante: *elle revient à dire que l'on peut se donner à la fois et indépendamment le lieu géométrique L de la moyenne μ de Y et son tenseur de variance V.* Nous noterons $\mathscr{P}(L, \{V\})$ une telle structure statistique sur E.

1.3. L'estimation au sens de Gauss–Markov. Soient L le sous-espace vectoriel de E engendré par $\{\mu_\theta: \theta \in \Theta\}$ et B une application linéaire de E dans F. La linéarité de l'espérance mathématique entraine clairement qu'une statistique linéaire U à valeurs dans F est un estimateur sans biais de $B\mu$ si et seulement si $B-U$ s'annule sur L. De la proposition 2.1 de [20], nous déduisons donc le théorème suivant dans lequel $L^\perp$ désigne l'orthogonal de L dans E^*:

THÉORÈME 1.1. *Pour toute application linéaire B de E dans F, l'ensemble des estimateurs linéaires sans biais de B est le sous-espace affine $\Lambda(B)$ de $\mathscr{L}(E, F)$ défini par*

$$\Lambda(B) = B + L^\perp \otimes F.$$

Ce théorème s'applique aussitôt au modèle de Gauss–Markov $\mathscr{P}(L, \{V\})$ sur E; autrement dit, la condition $L \subset \operatorname{Ker}(B-U)$ est nécessaire et suffisante pour qu'un estimateur linéaire U de B soit sans biais.

Nous appellerons *estimateur de Gauss–Markov de $B\mu$* (ou *de B*) tout élément minimum $\bar{B}$ de $\Lambda(B)$ pour le préordre $\preccurlyeq_1$, autrement dit tout élément $\bar{B}$ de $\Lambda(B)$, s'il en existe, défini par

$$(\forall \theta \in \Theta)\ (\forall U \in \Lambda(B))\ \bar{B} \circ V_\theta \circ {}^t\bar{B} \preccurlyeq_L U \circ V_\theta \circ {}^tU.$$

La recherche d'estimateurs de Gauss–Markov de B se ramène donc à la résolution de programmes quadratiques étudiés dans la section 3 de [20]. Nous retiendrons ici notamment que:

– la recherche d'estimateurs de Gauss–Markov de la moyenne μ est capitale, car pour tout estimateur de Gauss–Markov Q de μ, $B \circ Q$ est un estimateur de Gauss–Markov de $B\mu$ et tout estimateur de Gauss–Markov de $B\mu$ est $\mathscr{P}$-presque-sûrement de cette forme (conséquence du § 3.2 et du corollaire 3.2 de [20]);

— si F est mis en dualité séparante avec lui-même, tout estimateur de Gauss–Markov de $B\mu$ est un élément minimum de $\Lambda(B)$ pour le préordre $\preccurlyeq_2$ si et seulement si cette dualité est positive sur F, ce qui revient à supposer F euclidien (cf. théorème 3.6 de [20] et la conclusion qui suit).

Enfin, des théorèmes 3.4 et 3.5 de [20], on peut dégager à la fois des conditions de non vacuité de l'ensemble des estimateurs de Gauss–Markov de toute application linéaire B de E dans F et une caractérisation géométrique de la famille des estimateurs de Gauss–Markov de la moyenne μ (cas où $B = I_E$, application identique de E; se reporter à [21] pour une étude plus complète). Nous énoncerons ici la version statistique suivante de la conjonction des théorèmes 3.4 et 3.5 de [20], utile pour la suite (voir également [15], p. 273):

THÉORÈME 1.2. *Soit $\mathscr{P}(L, \{V\})$ un modèle de Gauss–Markov sur E, L désignant un sous-espace vectoriel de E. Pour qu'un endomorphisme Q de E soit un estimateur de Gauss–Markov de μ, il faut et il suffit que l'une ou l'autre des deux conditions équivalentes suivantes soit vérifiée*:

(i) *$L^{\perp}$ est inclus dans* $\operatorname{Ker}(Q \circ V)$;

(ii) *La restriction de Q au sous-espace $L + \operatorname{Im} V$ (qui est le plus petit sous-espace vectoriel de E de mesure l'unité pour toutes les probabilités mises dans le modèle) est le projecteur sur L parallèlement à $V(L^{\perp})$.*

1.4. L'optimalité et l'admissibilité en estimation linéaire. Il est immédiat que la recherche d'estimateurs linéaires optimaux ou admissibles d'une application linéaire B de E dans F se ramène à l'étude menée dans la section 4 de [20]. Lorsque nous parlerons d'admissibilité, nous supposerons implicitement que F est muni d'une structure euclidienne (se reporter aux remarques 1 et 2 du § 4.2 de [20]). Evidemment, l'optimalité et l'admissibilité s'entendent prises relativement à l'ensemble Σ (ou ce qui est équivalent, *relativement au cône convexe fermé engendré par* Σ) ce qui, sauf mention explicite du contraire, ne sera plus spécifié.

Nous noterons Ξ le sous-espace vectoriel de $\mathscr{L}_s(E^*, E) \times \mathscr{L}_s(E^*, E)$ engendré par Σ, puis $\varrho(M, N, U)$ l'extension de Σ à Ξ du tenseur de l'erreur quadratique d'un estimateur linéaire U d'une application linéaire B de E dans F, autrement dit:

$$(\forall (M, N) \in \Xi) \ \varrho(M, N, U) = U \circ M \circ {}^t U + (U - B) \circ N \circ {}^t(U - B).$$

Conformément à l'usage, nous nommerons *estimateur linéaire admissible de $B\mu$* (ou *de B*) tout estimateur admissible de B dans l'espace $\mathscr{L}(E, F)$.

Tous les résultats obtenus dans la section 4 de [20] s'appliquent bien entendu au contexte de l'estimation linéaire avec toutefois une légère simplification puisque l'ensemble Σ est maintenant constitué de couples (M, N) de tenseurs *positifs sur tout l'espace* E^* (pour chaque θ de Θ, $M = V_\theta$ et $N = \mu_\theta \otimes \mu_\theta$). Ils se transcrivent immédiatement en termes d'estimation linéaire. L'ensemble des estimateurs linéaires optimaux ou admissibles d'une application linéaire B possèdent de nombreuses propriétés communes. En particulier, nous retiendrons le résultat suivant, déduit du théorème 4.12 de [20] et extension significative d'un théorème de Olsen et al. (cf. [14], proposition 3.3):

THÉORÈME 1.3. *Soient Z un sous-espace vectoriel de E^* et Γ un sous-espace affine de $\mathscr{L}(E, F)$ de direction $Z \otimes W$. L'ensemble des estimateurs linéaires optimaux [resp. admissibles] dans Γ de chaque élément B de Γ n'est pas vide et forme une classe complète minimale.*

Evidemment, lorsque $Z = L^\perp$, ce théorème concerne l'ensemble des estimateurs linéaires *sans biais* optimaux [resp. admissibles] d'une application linéaire B (cf. théorème 1.1). Lorsque $Z = E^*$, il s'applique à l'ensemble des estimateurs linéaires optimaux [resp. admissibles] de B.

2. ESTIMATEURS LINÉAIRES ADMISSIBLES DE LA MOYENNE SUR UN MODÈLE GÉNÉRAL DE GAUSS–MARKOV

Les notations sont celles de la section 1 précédente. Dans toute la suite, nous supposons désormais que $(E, \mathscr{B}_E, \mathscr{P})$ est un modèle général de Gauss–Markov, autrement dit, que Σ est un produit cartésien de la forme

$$\Sigma = \{V\} \times \{u \otimes u:\ u \in L\},$$

où L est un sous-espace vectoriel de E et V un élément non nul de $\mathscr{L}_s^+(E^*, E)$.

Afin de donner une certaine unité à la suite de cet article (puisque nous chercherons notamment à caractériser l'ensemble des estimateurs linéaires admissibles de la moyenne μ), nous supposons que E et F sont *euclidiens* et identifiés, sauf mention explicite du contraire, à leurs duaux respectifs E^* et F^*, d'où l'invariance du choix de ces structures. Pour chaque sous-ensemble $\mathscr{X}$ de $\mathscr{L}_s(E) \times \mathscr{L}_s(E)$, nous noterons $[\mathscr{X}]$ le cône convexe fermé de l'espace-produit $\mathscr{L}_s(E) \times \mathscr{L}_s(E)$ engendré par $\mathscr{X}$, autrement dit: l'adhérence de l'ensemble des combinaisons linéaires à coefficients strictement positifs des éléments de $\mathscr{X}$.

Il en découle immédiatement que

$$\begin{aligned}[][\Sigma] &= \{(\lambda V, N):\ \lambda \geqslant 0,\ N \in \mathscr{L}_s^+(E),\ \operatorname{Im} N \subset L\} \\ &= \boldsymbol{R}^+ V \times \{N \in \mathscr{L}_s^+(E):\ \operatorname{Im} N \subset L\}.\end{aligned}$$

De même, le sous-espace vectoriel Ξ de $\mathscr{L}_s(E) \times \mathscr{L}_s(E)$ engendré par Σ est clairement le produit cartésien:

$$\Xi = \boldsymbol{R} V \times \{N \in \mathscr{L}_s(E):\ \operatorname{Im} N \subset L\}.$$

Enfin, le sous-espace $L + \operatorname{Im} V$ de E coïncide ici avec la somme, notée $[\Sigma]^\#$, des sous-espaces $\operatorname{Im}(M+N)$ quand (M, N) décrit $[\Sigma]$.

Nous allons chercher à caractériser l'ensemble des estimateurs linéaires admissibles (dans $\mathscr{L}(E, F)$) d'une application linéaire donnée B de E dans F. Nous poserons donc $\Gamma = \mathscr{L}(E, F)$ et utiliserons ainsi naturellement les notations des sections 3 et 4 de [20]. Soit $\mathscr{S}(B, \Gamma)$ le sous-espace vectoriel de Ξ formé des éléments singuliers (cf. § 4.2 de [20]). Comme ici $Z = E$, il est immédiat que l'on a

$$\mathscr{S}(B, \Gamma) = \{(M, N) \in \Xi : \ M + N = 0\},$$

et par conséquent que $\mathscr{S}(B, \Gamma)$ ne dépend pas de B, et que

$$[\Sigma] \cap \mathscr{S}(B, \Gamma) = \{0\}.$$

Nous noterons donc simplement $\mathscr{S}(B, \Gamma) = \mathscr{S}$.

2.1. Conditions nécessaires d'admissibilité en estimation linéaire de la moyenne sur un modèle de Gauss–Markov. La recherche de conditions nécessaires d'admissibilité en estimation linéaire sur un modèle général de Gauss–Markov repose sur la proposition nouvelle suivante, B désignant une fois pour toutes une application linéaire de E dans F:

PROPOSITION 2.1. *Supposons $L + \operatorname{Im} V = E$ et soit A un estimateur linéaire de B. Si A est admissible, alors:*

– ou il existe $(\lambda V, N)$ dans $[\Sigma]$, avec $\lambda > 0$, tel que A appartienne à $\Gamma_{\lambda V, N}(B)$;

– ou sinon, il existe un couple $(0, N_1)$ dans $[\Sigma]$, avec $N_1 \neq 0$, tel que A appartienne à $\Gamma_{0, N_1}(B)$, et un couple $(\lambda V, N_2)$ de $[\Sigma + \mathscr{S}_1]$ avec $\lambda > 0$, tel que A appartienne à $(\Gamma_1)_{\lambda V, N_2}(B)$, Γ_1 désignant l'ensemble $\Gamma_{0, N_1}(B)$ et $\mathscr{S}_1$ le sous-espace des éléments singuliers relatifs à Γ_1.

Lorsque V est bijectif, l'une ou l'autre de ces deux conditions est suffisante pour que A soit un estimateur linéaire admissible de B.

Démonstration. En effet, en raison du théorème 4.12 de [20], pour qu'un élément A de Γ soit un estimateur admissible de B, il faut et il suffit qu'il existe un couple non nul $(\lambda_1 V, N_1)$ de $[\Sigma]$ tel que A appartienne à $\Gamma_{\lambda_1 V, N_1}(B)$ et que A soit admissible pour B dans l'ensemble $\Gamma_{\lambda_1 V, N_1}(B)$.

Ou $\lambda_1 > 0$, et alors la première partie de la proposition est vérifiée, ou sinon $\lambda_1 = 0$, et alors A appartient au sous-espace affine

$$\Gamma_{0, N_1}(B) = B + (\operatorname{Ker} N_1) \otimes F \quad \text{avec } N_1 \neq 0.$$

Posons $\Gamma_1 = \Gamma_{0, N_1}(B)$ et $\mathscr{S}_1 = \mathscr{S}(B, \Gamma_1)$. Vérifions que tout couple $(\lambda V, N)$ du cône convexe $[\Sigma + \mathscr{S}_1] = \big[[\Xi] + \mathscr{S}_1\big]$ est tel que $\lambda \geqslant 0$ et que N soit positif sur $\operatorname{Ker} N_1$, et que, si $(\lambda V, N)$ appartient à $\mathscr{S}_1$, alors $\lambda = 0$. En effet, si $(\lambda V, N)$ est un couple de $[\Sigma + \mathscr{S}_1]$, la restriction de $\lambda V + N$ à $\operatorname{Ker} N_1$ est positive, et donc $N \succcurlyeq_L -\lambda V$ sur $\operatorname{Ker} N_1$. Or L contient $\operatorname{Im} N_1$ (car Σ engendre linéairement Ξ), et donc, par passage aux orthogonaux, $N \succcurlyeq_L -\lambda V$ sur $L^{\perp}$. Si $L = E$, la propriété est donc établie. Si L est strictement inclus dans E, on raisonne par l'absurde en supposant $\lambda < 0$: la restriction de $-\lambda V$ à $L^{\perp}$ serait alors strictement positive, et donc aussi celle de N, et par suite on aurait

$$L^{\perp} \cap \operatorname{Ker} N = \{0\} \quad \text{et} \quad L = E,$$

car $L^{\perp} \cap \operatorname{Ker} N = \{0\}$ équivaut à $L + \operatorname{Im} N = E$ et que $\operatorname{Im} N \subset L$. La contradiction en résulte vu que $L + \operatorname{Im} V = E$. Il s'ensuit que, si $(\lambda V, N)$ appartient à $\mathscr{S}_1$, alors $\lambda = 0$. En effet, $\mathscr{S}_1$ est inclus dans le cône convexe $[\Sigma + \mathscr{S}_1]$, et donc tout $(\lambda V, N)$ de $\mathscr{S}_1$ est tel que $\lambda \geqslant 0$ et que N soit positif sur $\operatorname{Ker} N_1$. D'autre part (cf. proposition 3.7 de [20]), tout $(\lambda V, N)$ de $\mathscr{S}_1$ vérifie

$$\operatorname{Ker} N_1 \subset \operatorname{Ker}(\lambda V + N);$$

autrement dit, $\lambda V = -N$ sur $\operatorname{Ker} N_1$, et donc $\lambda = 0$ (car V est positif sur tout l'espace E).

Si $\lambda_1 = 0$, l'admissibilité de A pour B dans $\Gamma_1 = \Gamma_{0,N_1}(B)$ est donc équivalente à l'existence d'un couple $(\lambda_2 V, N_2)$ non nul de $[\Sigma + \mathscr{S}_1] \backslash \mathscr{S}_1$ tel que A appartienne à l'ensemble $\Gamma_2 := (\Gamma_1)_{\lambda_2 V, N_2}(B)$ et que A soit admissible pour B dans Γ_2. A nouveau, λ_2 est positif ou nul, N_2 est positif sur $\operatorname{Ker} N_1$ et, si $\lambda_2 = 0$, tout $(\lambda V, N)$ de $\mathscr{S}_2 = \mathscr{S}(B, \Gamma_2)$ est tel $\lambda = 0$.

Or, en vertu de la proposition 4.13 de [20], une telle procédure ne comporte qu'un nombre fini d'étapes, ici toutes semblables étant donné la nature de l'ensemble Σ. Il n'est donc pas restrictif de supposer que l'étape ci-dessus est la dernière, ce qui entraîne (cf. proposition 4.13 de [20]) que

$$\operatorname{Rg}(\lambda_2 V + N_2)_{Z_1} = \operatorname{Max}\{\operatorname{Rg}(\lambda V + N)_{Z_1} \colon (\lambda V, N) \in [\Sigma + \mathscr{S}_1]\},$$

où $Z_1 = \operatorname{Ker} N_1$ et donc Σ est inclus dans $\mathscr{S}_2$ (cf. théorème 4.9 de [20] et la remarque qui suit). Il en résulte immédiatement que $\lambda_2 > 0$ car, si $\lambda_2 = 0$, chaque $(\lambda V, N)$ de $\mathscr{S}_2$ est tel que $\lambda = 0$, et Σ ne pourrait être contenu dans $\mathscr{S}_2$, ce qui termine la preuve de la première partie de cette proposition.

La deuxième partie est une conséquence immédiate du théorème de LaMotte (théorème 4.12 de [20]), les sous-espaces affines $\Gamma_{\lambda_1 V, N_1}(B)$ et $(\Gamma_1)_{\lambda_2 V, N_2}(B)$ étant réduits chacun à un point si V est bijectif et si λ_1 et λ_2 sont strictement positifs (cf. théorèmes 3.1 et 3.6 de [20]). ∎

L'ensemble des estimateurs linéaires admissibles de B coïncide avec l'ensemble des estimateurs linéaires admissibles de *chaque* élément C de la famille $\Lambda(B)$ des estimateurs linéaires sans biais de B (cf. proposition 4.4 de [20] et théorème 1.1 précédent). Il résulte alors immédiatement des théorème 4.15 et corollaire 3.3 de [20], et de la proposition 2.1 précédente, la proposition suivante réunissant un ensemble de conditions nécessaires d'admissibilité d'une application linéaire B définie sur un modèle de Gauss–Markov $\mathscr{P}(L, \{V\})$. On pourra rapprocher cette proposition du théorème 3.1 de [9] qui fournit (dans un cadre matriciel) un ensemble de conditions nécessaires d'admissibilité en estimation linéaire sur un modèle linéaire plus général:

PROPOSITION 2.2. *Pour qu'une statistique linéaire A à valeurs dans F soit un estimateur linéaire admissible de B, il faut que chacune des conditions suivantes soit vérifiée, C désignant un estimateur linéaire sans biais arbitraire de B*:

(i) $A(L+\operatorname{Im} V) = A(L)$ *et* $A(L) \subset B(L)$;

(ii) $(A-C)(L+\operatorname{Im} V) = \operatorname{Im}[(A-C)\circ V]$;

(iii) $(A-C)(L) = (A-C)(L \cap \operatorname{Im} V)$;

(iv) $L^{\perp} \subset \operatorname{Ker}(A \circ V)$;

(v) $A \circ V \circ {}^t C$ *est symétrique*;

(vi) $A \circ V \circ {}^t A \preccurlyeq_L A \circ V \circ {}^t C \preccurlyeq_L C \circ V \circ {}^t C$.

CONVENTION. Il est clair que la condition $A \circ V \circ {}^t A \preccurlyeq_L A \circ V \circ {}^t C$ exige que $A \circ V \circ {}^t C$ soit symétrique (car $A \circ V \circ {}^t A$ l'est toujours, cf. § 3.1 de [20]). C'est pourquoi désormais, lorsque nous écrirons la relation $A \circ V \circ {}^t A \preccurlyeq_L A \circ V \circ {}^t C$ (pour tous A et C de $\mathscr{L}(E, F)$), nous supposerons *implicitement* que $A \circ V \circ {}^t C$ est symétrique.

2.2. Caractérisations de l'admissibilité en estimation linéaire de la moyenne sur un modèle de Gauss–Markov. La recherche de conditions suffisantes d'admissibilité sur un modèle de Gauss–Markov repose sur le théorème suivant dû à Klonecki et Zontek (cf. théorème 4.1 de [9]) pour le cas univarié:

THÉORÈME 2.3. *Pour qu'un endomorphisme A de E soit un estimateur linéaire admissible de μ, il faut et il suffit que les relations suivantes soit vérifiées*:

(i) $A(L+\operatorname{Im} V) \subset L$;

(ii) $(I_E - A)(L+\operatorname{Im} V) = \operatorname{Im}[(I_E - A)\circ V]$;

(iii) $A \circ V \circ {}^t A \preccurlyeq_L A \circ V$.

Démonstration. Nous donnons une preuve fondée sur le théorème de LaMotte, basée sur la même idée directrice que celle du théorème 4.1 de [9], mais utilisant plus largement la structure linéaire de E (qui n'est pas nécessairement ici un espace $\boldsymbol{R}^n$). Pour cette preuve, nous pouvons supposer que $L+\operatorname{Im} V = E$ car les assertions (i) et (ii) ne concernent que la restriction de A et $I_E - A$ au sous-espace $L+\operatorname{Im} V$, tandis que l'assertion (iii) ne s'adresse qu'à la restriction de A au sous-espace $\operatorname{Im} V$. La nécessité de ces conditions vient d'être démontrée (proposition 2.2) en prenant $B = I_E$. Il reste donc à prouver leur suffisance.

Posons $T = {}^t(I_E - A)$. L'assertion (ii) équivaut alors à $\operatorname{Ker} {}^t T \oplus \operatorname{Im}(V \circ T) = E$ ($\oplus$ désignant la somme directe de deux sous-espaces), ou encore, d'après (iii), comme $V \circ T$ est symétrique, à la relation

$$\operatorname{Ker} {}^t T \oplus \operatorname{Im}({}^t T \circ V) = E.$$

Or (ii) équivaut aussi à $\operatorname{Im}({}^t T \circ V) = \operatorname{Im} {}^t T$, et donc, après passage aux orthogonaux, il est clair que T vérifie la relation

$$\operatorname{Ker} T \oplus \operatorname{Im} T = E \ (^1).$$

(1) C'est à partir de cette relation remarquable que nous utilisons plus largement que dans [9] (pp. 18–19) la structure linéaire de E.

Soit Π le projecteur de E sur $\operatorname{Im} T$ parallèlement à $\operatorname{Ker} T$. L'unique g-inverse réflexive T^- de T vérifiant $\Pi = T \circ T^-$ et $I_E - \Pi = T^- \circ T$ *commute avec* T. Il vient $\Pi \circ T = T \circ \Pi = T$, d'où l'on tire $A \circ {}^t(I_E - \Pi) = {}^t(I_E - \Pi)$, ce qui revient (la restriction de ${}^t(I_E - \Pi)$ à l'image de $I_E - \Pi$ étant injective) à

$$A \circ N_0 = N_0, \quad \text{où } N_0 = {}^t(I_E - \Pi) \circ (I_E - \Pi).$$

Par conséquent, d'après le corollaire 2.2 de [20], A décrit le sous-espace affine

$$\Phi = I_E + (\operatorname{Ker} N_0) \otimes E = I_E + (\operatorname{Im} T) \otimes E$$

de $\mathscr{L}(E)$. En raison des théorèmes 3.1 et 3.6 de [20], Φ n'est autre que l'ensemble des solutions du programme quadratique

$$\operatorname{Min} \operatorname{tr} [\varrho(0, N_0, U)]: \ U \in \mathscr{L}(E).$$

Donc, en vertu de la proposition 4.3 de [20], l'admissibilité de A dans $\mathscr{L}(E)$, en tant qu'estimateur linéaire de I_E, équivaut à la l'admissibilité de A *dans* Φ.

(a) Si $\Pi = 0$, alors N_0 est un tenseur bijectif et $\Phi = \{I_E\}$: dans ce cas, $A = I_E$ est un estimateur admissible de μ dans Φ, et donc aussi dans $\mathscr{L}(E)$.

(b) Supposons $\Pi \neq 0$ et posons $R = {}^t\Pi \circ A \circ V \circ \Pi$. D'après la condition (iii), R est un tenseur symétrique et positif sur E. De plus, comme $\Pi \circ T = T \circ \Pi$, nous avons

$$\Pi \circ {}^tA = {}^tA \circ \Pi \quad \text{et} \quad A \circ {}^t\Pi = {}^t\Pi \circ A,$$

ce qui entraîne $R = A \circ {}^t\Pi \circ V \circ \Pi$, et donc, d'après (i), l'inclusion $\operatorname{Im} R \subset L$. De plus,

$$\begin{aligned} R \circ T = R - R \circ {}^tA &= {}^t\Pi \circ A \circ V \circ \Pi - {}^t\Pi \circ A \circ V \circ \Pi \circ {}^tA \\ &= {}^t\Pi \circ A \circ V \circ \Pi - {}^t\Pi \circ A \circ V \circ {}^tA \circ \Pi \\ &= {}^t\Pi \circ (A \circ V - A \circ V \circ {}^tA) \circ \Pi \end{aligned}$$

et par suite, d'après (iii), $R \circ T$ est symétrique et positif sur E.

Posons $N_1 = R \circ T^-$. De $R = R \circ \Pi = R \circ T \circ T^-$, on tire

$${}^tN_1 = {}^t(T^-) \circ R = {}^t(T^-) \circ R \circ T \circ T^-,$$

et on en déduit que N_1 est symétrique et positif sur E. Comme, d'autre part, le sous-espace $\operatorname{Im} R$ est inclus dans L, nous avons $\operatorname{Im} N_1 \subset L$ et $(V, N_1) \in [\Sigma]$. Enfin, l'égalité $\Pi \circ T^- = T^-$ entraîne que, $A \circ V$ étant symétrique d'après (iii):

$$\begin{aligned} (I_E - A) \circ N_1 &= (I_E - A) \circ {}^t\Pi \circ A \circ V \circ T^- \\ &= {}^t(\Pi \circ T) \circ A \circ V \circ T^- = (I_E - A) \circ A \circ V \circ T^- \\ &= (I_E - A) \circ V \circ {}^tA \circ T^- = A \circ V \circ \Pi, \end{aligned}$$

car ${}^tA \circ T^- = (I_E - T) \circ T^- = T^- - \Pi$. Il s'ensuit aussitôt la relation

$$A \circ (V + N_1) \circ \Pi = N_1 \circ \Pi,$$

autrement dit, N_0 et Π ayant même image, que A appartient à $\Phi_{VN_1}(I_E)$ (cf. théorèmes 3.1 et 3.6 de [20]). Or, la restriction de V au sous-espace $\operatorname{Ker} N_0 = \operatorname{Im} \Pi$ est strictement positive [assertion (ii), vu que $L + \operatorname{Im} V = E$], et il en est aussi de même pour $V + N_1$. L'ensemble $\Phi_{VN_1}(I_E)$ est donc un singleton (cf. théorème 3.1 de [20]). L'admissibilité de A dans Φ est alors une conséquence du théorème 4.12 de [20]), ce qui achève la preuve. ∎

Remarque importante. Lorsque $L + \operatorname{Im} V = E$, la condition (ii) revient à dire que la restriction de V au sous-espace $\operatorname{Im} {}^t(I_E - A)$ est injective. Il en découle que l'on obtient un théorème équivalent en substituant à (ii) la condition

(ii)′ $(I_E - A)(L) \subset \operatorname{Im}[(I_E - A) \circ V]$.

Lorsque V est bijectif, la condition (ii) est trivialement vérifiée tandis que, si $A \circ V$ est symétrique, (iii) équivaut à $A^2 \circ V \preccurlyeq_L A \circ V$, autrement dit, à $A^2 \preccurlyeq_L A$ (V étant bijectif), ce qui revient à dire que le spectre de A est inclus dans $[0, 1]$, d'où le résultat suivant, nouveau sous cette forme et extension du corollaire 3.1 de [13] (p. 3034):

COROLLAIRE 2.4. *Lorsque V est bijectif, l'ensemble des estimateurs linéaires admissibles de μ est identique à l'ensemble des endomorphismes A de E d'image contenue dans L tels que $A \circ V$ soit symétrique et dont le spectre est inclus dans* $[0, 1]$.

Nous énoncerons également la

PROPOSITION 2.5. *Soit C un estimateur linéaire de B. Lorsque L est inclus dans* $\operatorname{Im} V$, *la conjonction des conditions*

(i) $L^\perp \subset \operatorname{Ker}(A \circ V)$, *et*

(ii) $A \circ V \circ {}^tA \preccurlyeq_L A \circ V \circ {}^tC$,

est nécessaire et suffisante pour qu'un estimateur linéaire A de B soit admissible.

Démonstration. Pour cette preuve, nous pouvons encore supposer que $L + \operatorname{Im} V = E$ sans perte de généralité. Il est alors clair que, si L est inclus dans $\operatorname{Im} V$, toutes les conditions de la proposition 2.2 se réduisent aux conditions de ce théorème, qui sont donc nécessaires pour que A un estimateur linéaire admissible de B. Vérifions qu'elles sont aussi suffisantes.

Posons $\mathscr{V} = A \circ V \circ {}^tC$ et $R = V \circ {}^tA \circ \mathscr{V}^\dagger \circ A$, où $\mathscr{V}^\dagger$ est le pseudo-inverse de Moore–Penrose de $\mathscr{V}$ (cf. 1.1). Comme d'après (ii) $\mathscr{V}$ est symétrique, on a

$$E = \operatorname{Ker} \mathscr{V} \oplus \operatorname{Im} \mathscr{V},$$

et il en découle que:

(a) $\operatorname{Im} R \subset L$ car, d'après (i), $L \subset (\mathscr{V} \circ {}^tA)$;

(b) $R \circ V = V \circ {}^tA \circ \mathscr{V}^\dagger \circ A \circ V$ est symétrique (car $\mathscr{V}^\dagger$ est symétrique);

(c) $R \circ V \circ {}^tR = V \circ {}^tA \circ \mathscr{V}^\dagger \circ (A \circ V \circ {}^tA) \circ \mathscr{V}^\dagger \circ A \circ V$, où d'après (ii),

$$R \circ V \circ {}^tR \preccurlyeq_L V \circ {}^tA \circ \mathscr{V}^\dagger \circ \mathscr{V} \circ \mathscr{V}^\dagger \circ A \circ V = V \circ {}^tA \circ \mathscr{V}^\dagger \circ A \circ V,$$

et par suite:

$$R \circ V \circ {}^t R \preccurlyeq_L R \circ V.$$

D'autre part, comme $L + \mathrm{Im}\, V = E$, l'hypothèse $L \subset \mathrm{Im}\, V$ implique que V est bijectif. La condition

$$\mathrm{Im}\,(I_E - R) = \mathrm{Im}\,[(I_E - R) \circ V]$$

est donc satisfaite, et le théorème 2.3 précédent montre que R est bien un estimateur admissible de μ.

Or, $\mathscr{V} \circ \mathscr{V}^\dagger$ a pour image $\mathrm{Im}\, \mathscr{V}$ et V est bijectif, donc

$$\mathrm{Im}\, A = \mathrm{Im}\,(A \circ V) = \mathrm{Im}\,(A \circ V \circ {}^t A) \subset \mathrm{Im}\, \mathscr{V}.$$

Le morphisme $\mathscr{V} = A \circ V \circ {}^t C$ étant symétrique (d'après (ii)), il s'ensuit que

$$C \circ R = \mathscr{V} \circ \mathscr{V}^\dagger \circ A = A,$$

autrement dit, en vertu du théorème 4.14 de [20] (stabilité de l'admissibilité par linéarité), que A est un estimateur admissible de C, et donc aussi de B. ∎

Nous obtenons alors le théorème caractéristique suivant, aménagement sensible du théorème 4.3 de [9] (p. 20):

THÉORÈME 2.6. *Soit C un estimateur linéaire sans biais arbitraire de B. Pour qu'un estimateur linéaire A de B soit admissible, il faut et il suffit que chacune des relations suivantes soit vérifiée*:

(i) $A(L + \mathrm{Im}\, V) \subset C(L)$;
(ii) $(A - C)(L + \mathrm{Im}\, V) = \mathrm{Im}\,[(A - C) \circ V]$;
(iii) $L^\perp \subset \mathrm{Ker}\,(A \circ V)$;
(iv) $A \circ V \circ {}^t A \preccurlyeq_L A \circ V \circ {}^t C$.

Démonstration. Vu la proposition 2.2, toutes ces conditions sont bien nécessaires pour que A soit un estimateur admissible de B. Il reste donc à pouver qu'elles sont suffisantes: pour cela *il suffit de prendre $C = B_*$ dans la famille $A(B)$ des estimateurs linéaires sans biais de B tel que* $\mathrm{Ker}\, B_* \subset L$. En effet, il existe bien un tel élément B_* dans $A(B)$ (cf. proposition 3.4 de [20]), et si ces conditions sont vérifiées pour $C = B_*$ et entraînent l'admissibilité de A, alors elles sont vérifiées pour chaque C de $A(B)$ (cf. proposition 2.2) ([2]).

Observons à nouveau que l'on ne restreint pas la généralité en supposant $L + \mathrm{Im}\, V = E$. En effet, les assertions (i) et (ii) de ce théorème ne concernent que les restrictions de A et B_* au sous-espace $L + \mathrm{Im}\, V$, tandis que les conditions (iii) et (iv) ne s'adressent qu'aux restrictions de ces applications au sous-espace $\mathrm{Im}\, V$ de E.

Posons

$$L_1 = L \cap \mathrm{Im}\, V \quad \text{et} \quad \Sigma^{(1)} = \{(V, u \otimes u) \colon\ u \in L_1\}.$$

([2]) Cette propriété constitue la clé de notre démonstration, démonstration que l'on pourra rapprocher de celle du théorème 4.3 de [9] (pp. 20–21) dans un cadre univarié.

Nous distinguerons l'admissibilité relativement aux ensembles Σ et $\Sigma^{(1)}$, nommées respectivement Σ-admissibilité et $\Sigma^{(1)}$-admissibilité (cf. définition 4.2 de [20]). Comme L contient L_1, (iii) entraîne la relation $L_1^\perp \subset \operatorname{Ker}(A \circ V)$, et les conditions (iii) et (iv) sont bien suffisantes, d'après la proposition 2.5 précédente, pour que A soit un estimateur $\Sigma^{(1)}$-admissible de B. Reprenant la preuve de cette proposition, il existe donc un endomorphisme R_1 de E qui est un estimateur $\Sigma^{(1)}$-admissible de μ vérifiant la relation $A = B_* \circ R_1$.

Il suffit de démontrer qu'il existe un estimateur linéaire Σ-admissible R de μ tel que $A = B_* \circ R$, ce qui prouvera (stabilité par linéarité de l'admissibilité; cf. théorème 4.14 de [20]) que A est Σ-admissible pour B_*, et donc pour B.

Comme $L + \operatorname{Im} V = E$ et que, pour tout C de $\Lambda(B)$, on a $B_*(L) = C(L)$, notons que la condition (i) entraîne

$$(\forall C \in \Lambda(B)) \quad C^{-1}(\operatorname{Im} A) \subset C^{-1}[C(L)],$$

où $C^{-1}[C(L)] = L + \operatorname{Ker} C$ et $\operatorname{Ker}(A - C) \subset C^{-1}[\operatorname{Im} A]$.

Il en découle $\operatorname{Ker}(A - B_*) \subset L + \operatorname{Ker} B_*$, avec (par définition) $\operatorname{Ker} B_* \subset L$, et par suite

$$\operatorname{Ker}(A - B_*) \subset L.$$

Sous l'hypothèse $L + \operatorname{Im} V = E$, (ii) équivaut à

$$\operatorname{Im}(A - B_*) = \operatorname{Im}[(A - B_*) \circ V].$$

Donc, d'après le théorème de factorisation des applications linéaires, il existe un endomorphisme X de E tel que $A - B_* = (A - B_*) \circ V \circ X$, autrement dit, tel que

$$V \circ X = \Pi + U,$$

où Π est un projecteur de E sur un supplémentaire de $\operatorname{Ker}(A - B_*)$ dans E et U un endomorphisme de E dont l'image est contenue dans $\operatorname{Ker}(A - B_*)$. Or, V étant positif sur E, nous avons

$$\operatorname{Im}(A - B_*) = \operatorname{Im}[(A - B_*) \circ V] \Leftrightarrow [\operatorname{Ker}(A - B_*)] \oplus \operatorname{Im}[V \circ {}^t(A - B_*)] = E,$$

et donc, puisque $\operatorname{Ker}(A - B_*)$ est inclus dans L, il n'est pas restrictif de supposer que $V \circ X$ est un projecteur de E, nécessairement sur un sous-espace vectoriel de E qui contient un supplémentaire de $\operatorname{Ker}(A - B_*)$ dans E: c'est le cas si U est un projecteur de E sur un sous-espace de $\operatorname{Ker}(A - B_*)$ vérifiant $\operatorname{Im} \Pi \subset \operatorname{Ker} U$. Cela revient donc à dire que l'on peut choisir X dans $\mathscr{L}(E)$ tel que $V \circ X$ soit un projecteur de E sur $\operatorname{Im} V$ ou, ce qui équivalent, que X soit une g-inverse de V telle que l'image de $I_E - V \circ X$ soit contenue dans L, car $I_E - V \circ X$ est un projecteur de E sur un sous-espace vectoriel de $\operatorname{Ker}(A - B_*)$.

Posons alors

$$R = (I_E - V \circ X) + R_1 \circ V \circ X = I_E + (R_1 \circ V - V) \circ X.$$

Comme R_1 est un estimateur $\Sigma^{(1)}$-admissible de μ [dans $\mathscr{L}(E)$], R_1 vérifie les conditions de la proposition 2.2, et donc:

(1) $\operatorname{Im} R \subset L$, car $\operatorname{Im} R_1 \subset L$ et $\operatorname{Im}(I_E - V \circ X) \subset L$;

(2) $R \circ V = (V - V \circ X \circ V) + R_1 \circ V \circ X \circ V = R_1 \circ V$,
et donc $R \circ V$ est symétrique;

(3) $I_E - R = V \circ X - R_1 \circ V \circ X = (I_E - R_1) \circ V \circ X$,
d'où $(I_E - R) \circ V = (I_E - R_1) \circ V$,

et par suite

$$\operatorname{Im}(I_E - R) = \operatorname{Im}(I_E - R) \circ V.$$

(4) $R \circ V \circ {}^t R = R \circ {}^t(R \circ V) = R \circ {}^t(R_1 \circ V) = R \circ V \circ {}^t R_1 = R_1 \circ V \circ {}^t R_1$,
d'où l'on tire que

$$R \circ V \circ {}^t R \preccurlyeq_L R \circ V.$$

Il suit (théorème 2.3) que R est un estimateur Σ-admissible de μ dans $\mathscr{L}(E)$. Enfin, des relations

$$B_* \circ R_1 = A \quad \text{et} \quad A - B_* = (A - B_*) \circ V \circ X,$$

il découle aussitôt:

$$\begin{aligned} B_* \circ R &= B_* \circ (I_E - V \circ X) + B_* \circ R_1 \circ V \circ X = B_* \circ (I_E - V \circ X) + A \circ V \circ X \\ &= B_* + (A - B_*) \circ V \circ X = A, \end{aligned}$$

ce qu'il fallait prouver. ∎

Il est parfois avantageux de substituer à ce théorème caractéristique le théorème suivant prouvé par Baksalary et Markiewicz (cf. [4], pp. 351–352) dans le cas univarié à l'aide d'une approche de l'admissibilité tout à fait différente de la nôtre ([3]). Il nous semble important et utile d'en donner une démonstration issue du théorème 2.6 précédent:

THÉORÈME 2.7. *Pour qu'un estimateur linéaire A de B soit admissible, il faut et il suffit que chacune des trois conditions suivantes soit vérifiée, C désignant un estimateur linéaire sans biais arbitraire de B:*

(i) $(A - C)(L) = (A - C)(L \cap \operatorname{Im} V)$;

(ii) $L^\perp \subset \operatorname{Ker}(A \circ V)$;

(iii) $A \circ V \circ {}^t A \preccurlyeq_L A \circ V \circ {}^t C$.

Démonstration. La nécessité de ces conditions provient encore de la proposition 2.2. Il reste donc à établir qu'elles sont suffisantes pour que A soit un estimateur admissible d'un élément arbitraire C de l'ensemble $\mathcal{A}(B)$ des

([3]) Comme nous l'avons précisé dans l'introduction de cet article, l'approche de Baksalary et Markiewicz se situe sur une lignée de travaux ouverte par Rao dans [17]. Mais une telle approche n'est pas directe et repose sur des techniques exclusivement matricielles développées dans [13].

estimateurs linéaires sans biais de B. On peut encore supposer que $L+\operatorname{Im} V = E$, ce qui, de manière évidente dans la formulation de ce théorème, n'affaiblit pas la généralité. L'assertion (iii) entraîne (V étant positif sur E et $A \circ V \circ {}^tC$ symétrique) que

$$\operatorname{Ker}(A \circ V \circ {}^tC) \subset \operatorname{Ker}(A \circ V \circ {}^tA).$$

Par passage aux orthogonaux, il vient donc

$$\operatorname{Im}(A \circ V) = \operatorname{Im}(A \circ V \circ {}^tA) \subset \operatorname{Im}(A \circ V \circ {}^tC),$$

et par suite, d'après (ii), on a

$$\operatorname{Im}(A \circ V) \subset C(L).$$

D'autre part (vérification immédiate), la relation

$$(A-C)(L) = (A-C)(L \cap \operatorname{Im} V)$$

équivaut à

$$L \subset (L \cap \operatorname{Im} V) + \operatorname{Ker}(A-C).$$

Posant $K = \operatorname{Ker}(A-C)$, on en déduit aussitôt (comme $L \cap \operatorname{Im} V \subset L$) que

$$L \subset [K+(L \cap \operatorname{Im} V)] \cap L = K \cap L + L \cap \operatorname{Im} V.$$

Il suit alors

$$A(L) \subset A(K \cap L) + A(L \cap \operatorname{Im} V),$$

où $A(K \cap L) = \{Au\colon\ Au = Cu,\ u \in L\} \subset C(L)$. Comme

$$A(L \cap \operatorname{Im} V) \subset \operatorname{Im}(A \cap V) \quad \text{et} \quad \operatorname{Im}(A \circ V) \subset C(L),$$

on obtient $A(L) \subset C(L)$. D'autre part, l'hypothèse $L+\operatorname{Im} V = E$ entraîne la relation

$$\operatorname{Im} A = A(L) + \operatorname{Im}(A \circ V),$$

dont on en déduit

$$\operatorname{Im} A \subset C(L).$$

Enfin, appliquant $A-C$ à chacun des membres de la relation $E = L+\operatorname{Im} V$, on obtient

$$\begin{aligned} \operatorname{Im}(A-C) &= (A-C)(L) + \operatorname{Im}[(A-C) \circ V] \\ &= (A-C)(L \cap \operatorname{Im} V) + \operatorname{Im}[(A-C) \circ V], \end{aligned}$$

dont il résulte aussitôt que

$$\operatorname{Im}(A-C) \subset \operatorname{Im}[(A-C) \circ V].$$

Sous l'hypothèse $L+\operatorname{Im} V = E$, toutes les conditions du théorème 2.7 précédent sont ainsi satisfaites et la démonstration achevée. ∎

Remarques importantes. 1. La condition (ii) de ce théorème (que l'on retrouve également dans le théorème 2.6) est remarquable: en vertu du théorème 1.2, elle signifie que *tout estimateur linéaire admissible A d'une application linéaire B est un estimateur de Gauss–Markov de A* (*autrement dit, de sa moyenne $A\mu$*).

2. Dans la preuve du théorème 2.6, nous avons établi que, si $L+\mathrm{Im}\, V = E$, alors pour qu'un estimateur linéaire A de B soit admissible, il est nécessaire que A soit de la forme $A = B_* \circ R$, où B_* désigne un estimateur linéaire sans biais de B de noyau inclus dans L et R un estimateur admissible de μ. Or $B-B_*$ s'annule sur L (théorème 1.1) et $\mathrm{Im}\, R$ est inclus dans L (proposition 2.2), d'où

$$(B-B_*)\circ R = 0 \quad \text{et} \quad A = B\circ R.$$

Lorsque $L+\mathrm{Im}\, V$ est strictement contenu dans E, on n'obtient la même relation qu'avec les restrictions de A et R au sous-espace $L+\mathrm{Im}\, V$.

Inversement (stabilité par linéarité de l'admissibilité), pour tout estimateur linéaire admissible U de μ, $B\circ U$ est un estimateur linéaire admissible de B. Or (cf. théorèmes 2.6 ou 2.7), si (A, A') désigne un couple d'estimateurs linéaires de B dont les restrictions à $L+\mathrm{Im}\, V$ sont égales, l'admissibilité de A équivaut à l'admissibilité de A'. Nous en déduisons le théorème caractéristique suivant qui indique comment on peut, *sur un modèle de Gauss–Markov*, reconstituer l'ensemble des estimateurs linéaires admissibles de toute transformation linéaire de μ à partir de l'ensemble des estimateurs linéaires admissibles de μ:

THÉORÈME 2.8. *Pour qu'un estimateur linéaire A de $B\mu$ soit admissible, il est nécessaire et suffisant qu'il existe un estimateur linéaire admissible U de μ tel que les restrictions de A et $B\circ U$ au sous-espace $L+\mathrm{Im}\, V$ soient égales.*

Tout estimateur linéaire admissible A de B est donc $\mathscr{P}$-presque-sûrement égal à $B\circ U$, où U est un estimateur linéaire admissible de μ.

2.3. Caractérisation variationnelle de l'admissibilité en estimation linéaire sur un modèle de Gauss–Markov. Le corollaire 3.2 de [20] conduit à une caractérisation variationnelle évidente de l'estimation de Gauss–Markov: pour qu'une statistique linéaire $\bar{B}$ à valeurs dans F soit un estimateur de Gauss–Markov de B, il faut et il suffit que, pour tout v^* de F^*, la forme linéaire $v^*\circ\bar{B}$ soit un estimateur de Gauss–Markov de $v^*\circ B$ dans E^*. Le théorème 2.7 permet d'obtenir une caractérisation variationnelle voisine de l'admissibilité en estimation linéaire sur un modèle de Gauss–Markov. Nous énoncerons le résultat suivant à la fois plus général et plus précis que le corollaire 4.4 de [9] (p. 21) ou le théorème 3.4 de [17] (p. 1029):

PROPOSITION 2.9. *Pour qu'une statistique linéaire A à valeurs dans F soit un estimateur linéaire admissible de B, il faut et il suffit que $A\circ V\circ {}^tB$ soit un endomorphisme symétrique de E et que, pour tout v^* de F^*, $v^*\circ A$ soit un estimateur admissible de $v^*\circ B$ dans E^*.*

Démonstration. La condition nécessaire est une conséquence directe de la proposition 2.2 (pour ce qui est de la symétrie de $A\circ V\circ {}^tB$) et de la stabilité par linéarité de l'admissibilité. La réciproque s'obtient en remarquant que si, pour tout v^* de F, $v^* \circ A$ est admissible pour $v^* \circ B$ dans E^* (identifié à E), la condition (iii) du théorème 2.7 entraîne

$$(\forall v \in F) \quad \langle V({}^tAv), {}^tAv \rangle \leqslant \langle V({}^tBv), {}^tBv \rangle,$$

autrement dit, que $A \circ V \circ {}^tA \preccurlyeq_L A \circ V \circ {}^tB$. Les autres conditions du théorème 2.7 se démontrent directement à l'aide de la propriété élémentaire suivante, dont la vérification (aisée) est immédiate: pour tout couple (E_1, E_2) de sous-espaces vectoriels de E, l'inclusion $E_1 \subset E_2$ équivaut à $u^*(E_1) \subset u^*(E_2)$ pour tout u^* de E^*. ■

2.4. Admissibilité en estimation linéaire et estimation de Gauss–Markov. Des théorèmes 2.6 ou 2.7, il découle immédiatement le

THÉORÈME 2.10. *Pour qu'un estimateur linéaire sans biais A de B soit un estimateur linéaire admissible de B, il faut et il suffit que A soit un estimateur de Gauss–Markov de B.*

En effet, l'assertion (i) du théorème 2.7 est trivialement vérifiée pour chaque estimateur linéaire sans biais C de B (cf. théorème 1.1), tandis que, si A est un estimateur linéaire sans biais de B, la condition (ii) du théorème 2.7 équivaut à dire que A est un estimateur de Gauss–Markov de B (cf. théorème 1.2), l'assertion (iii) étant alors satisfaite (se reporter au corollaire 3.3 de [20]). Chaque estimateur de Gauss–Markov de B est donc un estimateur linéaire admissible de B (i.e. dans $\mathscr{L}(E, F)$).

Inversement, tout estimateur linéaire admissible A de B est un estimateur de Gauss–Markov de son espérance mathématique $A\mu$ (cf. condition (ii) du théorème 2.7 et théorème 1.2). Lorsque B est l'application identique de E, l'étude complète de la réciproque du théorème précédent conduit alors au théorème suivant:

THÉORÈME 2.11. *Pour que l'ensemble des estimateurs linéaires admissibles de μ coïncide avec l'ensemble de ses estimateurs de Gauss–Markov, il faut et il suffit que*

$$L \cap \operatorname{Im} V = \{0\}.$$

Ce résultat remarquable est dû à Baksalary et Mathew (cf. [7], p. 61) lorsque E est un espace $\boldsymbol{R}^n$. Nous en donnons une preuve géométrique simple mettant on oeuvre la structure algébrique induite sur E par le tenseur V.

Démonstration. Chaque estimateur de Gauss–Markov de μ étant un estimateur linéaire admissible de μ, il suffit de prouver que cette condition équivaut à dire que tout estimateur linéaire admissible de μ est un estimateur de Gauss–Markov de μ.

Si $L \cap \operatorname{Im} V = \{0\}$, l'assertion (i) du théorème 2.7 indique que tout estimateur linéaire admissible de μ est nécessairement un estimateur linéaire sans biais de μ; le théorème 2.10 précédent permet aussitôt de conclure.

Inversement, si tout estimateur linéaire admissible de μ est un estimateur de Gauss–Markov de μ, alors, en vertu du théorème 4.9 de [20], N désignant un tenseur symétrique et positif sur E et d'image L, tout élément Q de $\Gamma_{VN}(I_E)$, avec ici $\Gamma = \mathscr{L}(E)$, est un estimateur de Gauss–Markov de μ, autrement dit, d'après le théorème 1.2, un endomorphisme Q de E dont la restriction Q' au sous-espace $L + \operatorname{Im} V$ est le projecteur sur L parallèlement à $V(L^\perp)$. Or, d'après les théorèmes 3.1 et 3.6 de [20], chaque élément Q de $\Gamma_{VN}(I_E)$ vérifie la relation $Q \circ (V+N) = N$, c'est-à-dire

$$Q \circ V = (I_E - Q) \circ N,$$

où, puisque Q' est le projecteur sur L de direction $V(L^\perp)$,

$$(I_E - Q) \circ N = 0.$$

Il s'ensuit que $Q \circ V = 0$, autrement dit, que $\operatorname{Im} V$ est inclus dans le noyau $\operatorname{Ker} Q' = V(L^\perp)$ de Q'. On en déduit que $\operatorname{Im} V = V(L^\perp)$, d'où il découle que $L \cap \operatorname{Im} V = \{0\}$, car $V(L^\perp)$ est un supplémentaire de L dans $L + \operatorname{Im} V$, ce qui achève la preuve. ∎

Interprétation géométrique. Désignons par L_2 un supplémentaire dans L du sous-espace $L_1 = L \cap \operatorname{Im} V$ dans L, et respectivement par μ_1 et μ_2 les composantes de μ sur L_1 et L_2. La statistique identique Y de E se décompose $\mathscr{P}$-presque-sûrement en la somme de deux vecteurs aléatoires Y_1 et Y_2 respectivement à valeurs dans $\operatorname{Im} V$ et L_2, d'espérances mathématiques μ_1 et μ_2. Or le vecteur $Y_2 - \mu_2$ est $\mathscr{P}$-presque-sûrement nul, et donc dire que L_1 est réduit à $\{0\}$ revient à dire que μ_1 est nulle (donc entièrement connue): il n'y a véritablement plus, à proprement parler, de problème inférentiel, car Y_2 estime sans biais et avec un tenseur de variance nul l'autre composante μ_2 de μ.

2.5. L'admissibilité d'estimateurs affines d'une application linéaire définie sur une structure de Gauss–Markov. Nous désignons par $\mathscr{A}(E, F)$ l'ensemble des applications affines de E dans F. La structure de l'ensemble des estimateurs *affines* admissibles [dans l'ensemble $\mathscr{A}(E, F)$] d'une application *linéaire* B de E dans F se déduit de celle des estimateurs linéaires admissibles de B selon l'énoncé suivant:

THÉORÈME 2.12. *L'admissibilité d'un estimateur affine $A' = t_a \circ A$ de B, de partie linéaire A et de translation associée t_a, équivaut à la conjonction des deux conditions suivantes*:

(i) *a appartient au sous-espace $(A-B)(L)$;*

(ii) *A est un estimateur linéaire admissible de B.*

Ce résultat, conjecturé par Rao (cf. corollaire 3.2 de [17]), a été effectivement démontré par Baksalary et Markiewicz (cf. théorème 1 de [2]) dans le cas univarié à l'aide d'arguments matriciels. Nous en donnons ici une preuve basée sur le langage fonctionnel lorsque E est un espace euclidien quelconque.

Démonstration. Pour tout $\theta = (\mu, V)$ de l'espace paramétrique Θ, identifié ici au produit $L\times\{V\}$, le tenseur de l'erreur quadratique d'un estimateur affine $A' = t_a\circ A$ de B s'écrit:

$$\varrho(\theta, t_a\circ A) = A\circ V\circ{}^tA+[(A-B)\mu+a]\otimes[(A-B)\mu+a].$$

Supposons que $t_a\circ A$ soit un estimateur admissible de B. Par contraposition, si a n'appartenait pas au sous-espace $(A-B)(L)$, a s'écrirait $a = a_1+a_2$, avec a_1 et a_2 éléments de $(A-B)(L)$ et de $[(A-B)(L)]^\perp$ respectivement, et $a_2 \neq 0$: il en résulterait clairement que

$$(\forall\theta\in\Theta)\ \operatorname{tr}[\varrho(\theta, t_a\circ A)-\varrho(\theta, t_{a_1}\circ A)] = \|a_2\|^2 > 0,$$

ce qui contredirait l'admissibilité de A' dans $\mathscr{A}(E, F)$. La condition (i) est donc nécessaire. Maintenant, si (ii) n'était pas vraie, alors il existerait un estimateur linéaire U de B strictement meilleur que A au sens où on aurait

$$(\forall m\in L)\ \operatorname{tr}[U\circ V\circ{}^tU]+\|(U-B)m\|^2 \leqslant \operatorname{tr}[A\circ V\circ{}^tA]+\|(A-B)m\|^2,$$

cette inégalité étant stricte pour au moins un élément m_* de L. La condition (i) étant satisfaite, tout m de L s'écrit $m=\mu+\alpha$, où est un antécédent *fixé* de a dans L par $A-B$ et μ un élément de L.

Posons $u=(U-B)\alpha$. Nous aurions alors

$$(U-B)m=(U-B)\mu+u \quad \text{et} \quad (A-B)m=(A-B)\mu+a,$$

et il en découlerait que $t_u\circ U$ serait strictement meilleur que $t_a\circ A$ dans $\mathscr{A}(E, F)$, ce qui mettrait en défaut l'admissibilité de $t_a\circ A$. La nécessité des conditions (i) et (ii) est ainsi prouvée.

La réciproque s'établit également par contraposition. Supposons les conditions (i) et (ii) vérifiées et $t_a\circ A$ non admissible. Il existerait alors un estimateur affine $t_u\circ U$ de B, de partie linéaire U, tel que

$$(\forall\mu\in L)\ \operatorname{tr}[U\circ V\circ{}^tU]+\|(U-B)\mu+u\|^2 \leqslant \operatorname{tr}[A\circ V\circ{}^tA]+\|(A-B)\mu+a\|^2,$$

cette inégalité étant stricte pour au moins un élément μ_* de L.

Soit α un antécédent *fixé* de a par $A-B$ appartenant à L. Posons $v=u+(U-B)\alpha$ et, pour chaque μ de L, $m=\mu+\alpha$. L'inégalité précédente devient alors

$$(\forall m\in L)\ \operatorname{tr}[U\circ V\circ{}^tU]+\|(U-B)m+v\|^2 \leqslant \operatorname{tr}[A\circ V\circ{}^tA]+\|(A-B)m\|^2,$$

et nous déduirions que

$$(\forall m\in L)\ \|(U-B)m+v\|^2 \leqslant \|(A-B)m\|^2.$$

Il en résulterait que $v=0$ (en prenant $m=0$), et A ne serait donc pas un estimateur linéaire admissible de B, ce qui contredirait (ii). ∎

Enfin, ce théorème 2.12 s'étend immédiatement au cas où L est un sous-espace affine de E que nous noterons L'. En effet, L désignant maintenant la direction de L', L' est l'ensemble des éléments $\mu = m+d$ de E, d décrivant L et m *fixé* choisi arbitrairement dans L'. Pour chaque $\theta = (d, V)$ de l'espace paramétrique Θ, identifié au produit cartésien $L \times \{V\}$, le tenseur d'erreur quadratique d'un estimateur affine $A' = t_a \circ A$ de B s'écrit:

$$\varrho(\theta, t_a \circ A) = A \circ V \circ {}^t\!A + [(A-B)d + (A-B)m + a] \otimes [(A-B)d + (A-B)m + a].$$

Le théorème précédent conduit à dire que pour que A' soit admissible, il faut et il suffit que $(A-B)m + a$ appartienne au sous-espace $(A-B)(L)$, autrement dit que

$$a \in (B-A)(L'),$$

et que A soit un estimateur linéaire admissible de B relativement au modèle $\mathscr{P}(L, \{V\})$. Par conséquent, a ne peut être nul que si L' est un sous-espace vectoriel, d'où le

COROLLAIRE 2.13. *Soit L' un sous-espace affine de E. Pour qu'une application linéaire B admette un estimateur linéaire admissible relativement au modèle $\mathscr{P}(L', \{V\})$, il faut et il suffit que L' soit un espace vectoriel de E. Sinon, un estimateur affine $A' = t_a \circ A$ de B, de partie linéaire A, est admissible, si et seulement si:*
(i) *a appartient au sous-espace $(B-A)(L')$, et*
(ii) *A est un estimateur linéaire admissible de B relativement au modèle linéaire $\mathscr{P}(L, \{V\})$, L désignant la direction de L'.*

Evidemment, si L' est un sous-espace vectoriel de E, alors

$$(B-A)(L') = (A-B)(L')$$

et on retrouve le théorème 2.12.

3. ESTIMATEURS LINÉAIRES ADMISSIBLES DU PARAMÈTRE DE RÉGRESSION

Dans toute cette section, nous conservons les hypothèses et notations de la section précédente, mais supposons en outre L est l'image d'une application linéaire T définie sur un espace euclidien H. Nous noterons $\mathscr{P}(Tb, \{V\})$ le modèle de régression linéaire correspondant, éventuellement multivarié, b désignant le paramètre de régression dans l'espace H.

Nous dirons qu'une application linéaire R définie sur H est *identifiable* si, pour tout couple (b_1, b_2) de $H \times H$ tel que $Tb_1 = Tb_2$, on a $Rb_1 = Rb_2$. Il équivaut à dire que $\operatorname{Ker} T \subset \operatorname{Ker} R$, ou encore que R admet un estimateur linéaire sans biais (cf. [1]).

Pour toute application linéaire R de H dans un espace euclidien F, le tenseur d'erreur quadratique de tout estimateur linéaire U de R s'écrit, pour chaque $\theta = (b, V)$ de $H \times \{V\}$:

$$\begin{aligned}\varrho(\theta, U) &= E_\theta[(U - Rb) \otimes (U - Rb)] \\ &= U \circ V \circ {}^tU + (U \circ T - R) \circ (b \otimes b) \circ {}^t(U \circ T - R).\end{aligned}$$

Par conséquent, si R est identifiable, le tenseur d'erreur quadratique d'un estimateur linéaire U de R n'est autre que le tenseur d'erreur quadratique de tout estimateur linéaire U d'une application linéaire arbitraire B de E dans F telle que $R = B \circ T$, car

$$\begin{aligned}(\forall b \in H) \quad \varrho(\theta, U) &= U \circ V \circ {}^tU + (U \circ T - B \circ T) \circ (b \otimes b) \circ {}^t(U \circ T - B \circ T) \\ &= U \circ V \circ {}^tU + (U - B) \circ (\mu \otimes \mu) \circ {}^t(U - B).\end{aligned}$$

Prolongeant naturellement la définition d'un estimateur linéaire admissible d'une application linéaire B définie sur E, nous nommerons *estimateur linéaire admissible d'une application linéaire R de H dans F tout estimateur admissible de R (ou Rb) dans $\mathscr{L}(E, F)$*. Il en découle clairement que l'ensemble des estimateurs linéaires admissibles d'une application linéaire identifiable R coïncide avec l'ensemble des estimateurs linéaires admissibles d'un estimateur linéaire sans biais arbitraire de R. Tous les résultats obtenus dans la section 2 s'appliquent donc immédiatement sachant que l'on a ici $C(L) = \operatorname{Im} R$ *pour chaque estimateur linéaire sans biais C de R.*

3.1. Caractérisation des estimateurs linéaires admissibles de b. Lorsque le paramètre b est identifiable (i.e. T injective), la recherche d'estimateurs linéaires ou affines admissibles du paramètre b est une question importante. Cela revient à prendre $R = I_H$ et C arbitraire dans l'ensemble des inverses à gauche T^- de l'application T.

Sous une telle hypothèse, si $T \circ A$ est un estimateur linéaire admissible de μ, alors (stabilité par linéarité de l'admissibilité, cf. théorème 4.14 de [20]) $T^- \circ T \circ A = A$ est un estimateur linéaire admissible de b. Usant de la même propriété, la réciproque est immédiate, ce qui conduit à la caractérisation suivante:

THÉORÈME 3.1. *Lorsque T est injective, pour qu'un estimateur linéaire A de b soit admissible, il faut et il suffit que $T \circ A$ soit un estimateur linéaire admissible de μ.*

Il en résulte immédiatement une série de caractérisations de l'ensemble des estimateurs linéaires admissibles de b issues de la section 2. Toutefois, T étant injective et le sous-espace $\operatorname{Im}(V \circ {}^tA)$ contenu dans $\operatorname{Im} T$ (cf. théorème 2.6 ou 2.7), nous avons, pour chaque inverse à gauche C de T, l'équivalence

$$(T \circ A) \circ V \circ {}^t(T \circ A) \preccurlyeq_L T \circ A \circ V \Leftrightarrow A \circ V \circ {}^tA \preccurlyeq_L A \circ V \circ {}^tC.$$

L'implication de droite à gauche est triviale. Inversement, de $\operatorname{Im}(V \circ {}^tA) \subset \operatorname{Im} T$, il suit (par passage aux orthogonaux) $A \circ V = A \circ V \circ \Pi$, où Π désigne le projecteur orthogonal de E sur L. Prenant $C = T^\dagger$ (pseudo-inverse de Moore–Penrose de T), nous pouvons écrire

$$T \circ A \circ V \circ {}^tC \circ {}^tT = T \circ A \circ V \circ {}^t(T \circ C) = T \circ A \circ V \circ \Pi = T \circ A \circ V,$$

et l'implication opposée de gauche à droite en résulte. Enfin, la relation

$$(T \circ A) \circ V \circ {}^t(T \circ A) \preccurlyeq_L T \circ A \circ V$$

entraîne la relation $\operatorname{Im}(V \circ {}^t\!A) \subset \operatorname{Im} T$. En effet, $T \circ A \circ V$ est alors symétrique et comme $C = T^\dagger$ est une inverse à droite de ${}^tT$ (car T est injective), nous avons

$$T \circ A \circ V \circ C = V \circ {}^t\!A \circ {}^tT \circ C = V \circ {}^t\!A,$$

dont il découle aussitôt l'inclusion cherchée. Nous obtenons donc le théorème suivant démontré par Baksalary et Markiewicz (lorsque E est un espace $\boldsymbol{R}^n$) à l'aide d'une méthode tout à fait différente (cf. [4], p. 358):

THÉORÈME 3.2. *Lorsque T est injective, l'ensemble des estimateurs linéaires admissibles de b est identique à l'ensemble des statistiques linéaires A à valeurs dans H vérifiant les deux conditions suivantes*:
(a) $(A-C)(L) = (A-C)(L \cap \operatorname{Im} V)$;
(b) $(T \circ A) \circ V \circ {}^t(T \circ A) \preccurlyeq_L T \circ A \circ V$.

Usant maintenant du corollaire 2.4, nous obtenons le résultat suivant dû à Baksalary et Markiewicz par une autre voie (cf. corollaire 1 de [2]):

COROLLAIRE 3.3. *Lorsque T est injective et V bijectif, l'ensemble des estimateurs linéaires admissibles de b se résume à l'ensemble des endomorphismes A de H tels que $T \circ A \circ V$ soit symétrique et tels que le spectre de $A \circ T$ soit contenu dans* $[0, 1]$.

3.2. La conjecture de C.R. Rao. Selon l'usage, nous appellerons *estimateur des moindres-carrés généralisés de b* toute statistique linéaire $\mathscr{C}$ à valeurs dans H telle que, pour tout y de E, $\hat{b} = \mathscr{C}y$ soit une solution de l'équation normale généralisée

$${}^tT \circ W^- \circ T\hat{b} = {}^tT \circ W^- y,$$

où W^- est une inverse généralisée arbitraire d'un tenseur symétrique W sur E^* (identifié à E) vérifiant les deux conditions:

$$L + \operatorname{Im} V = \operatorname{Im} W \quad \text{et} \quad W(L^\perp) = V(L^\perp). \tag{1}$$

Il est bien connu (cf. [16] ou [19] pour un traitement algébrique plus général de cette question) que $W = V + T \circ {}^tT$ satisfait ces deux conditions et que l'ensemble des estimateurs des moindres-carrés généralisés de b associé à une inverse généralisée W^- de W n'est réduit à un singleton $\{\mathscr{C}\}$ que si T est injective, et qu'alors

$$\mathscr{C} = ({}^tT \circ W^- \circ T)^{-1} \circ {}^tT \circ W^-.$$

Si T est injective, toute application linéaire R définie sur H est identifiable et chaque estimateur de Gauss–Markov $\bar{R}$ de R d'image $\operatorname{Im} R$ s'écrit $\bar{R} = R \circ \mathscr{C}$ (principe de substitution de Gauss, cf. [16], [19] ou [21] pour un traitement fonctionnel de cette question).

Une autre caractérisation de la structure de la famille des estimateurs linéaires admissibles A d'une application linéaire identifiable R résulte du fait que A est un estimateur de Gauss–Markov de lui-même (ou de sa moyenne $A\mu$, cf. remarque 1 suivant le théorème 2.7). Par conséquent, d'après le principe de substitution de Gauss, si A est un estimateur linéaire admissible de b, alors il existe un estimateur des moindres-carrés généralisés $\mathscr{C}$ de b, associé à une inverse généralisée W^- d'un tenseur symétrique W sur E vérifiant les conditions (1), tel que

$$A = A \circ T \circ \mathscr{C}.$$

Il revient à dire que $\bar{A} = A \circ T$ est un estimateur linéaire admissible de l'application identique de H relativement à la structure induite par $\mathscr{C}$ sur H. En effet. le tenseur d'erreur quadratique de $A = \bar{A} \circ \mathscr{C}$ s'écrit, si $\theta = (b, V)$:

$$\varrho(\theta, U) = \bar{A} \circ \mathscr{C} \circ V \circ {}^t\mathscr{C} \circ {}^t\bar{A} + (\bar{A} - I_H) \circ (b \otimes b) \circ {}^t(\bar{A} - I_H),$$

et $\mathscr{C}(L) = H$. Relevons que, si L est inclus dans $\operatorname{Im} V$, alors le tenseur de variance $\mathscr{C} \circ V \circ {}^t\mathscr{C}$ de $\mathscr{C}$ est bijectif et vaut $({}^tT \circ V^- \circ T)^{-1}$, V^- désignant une inverse généralisée *arbitraire* de V.

Posons alors $\mathscr{W} = {}^tT \circ V^- \circ T$. La structure statistique induite par $\mathscr{C}$ est le modèle de Gauss–Markov $\mathscr{P}(H, \{\mathscr{W}^{-1}\})$ sur H, et dire que A est un estimateur linéaire admissible de b revient à dire que $\bar{A}$ est un estimateur linéaire admissible de l'application identique de H relativement à la structure $\mathscr{P}(H, \{\mathscr{W}^{-1}\})$, autrement dit, que $\bar{A}$ vérifie la *seule* condition (iii) du théorème 2.3, à savoir:

$$\bar{A} \circ \mathscr{W}^{-1} \circ {}^t\bar{A} \preccurlyeq_L \bar{A} \circ \mathscr{W}^{-1}.$$

Nous énoncerons donc le théorème caractéristique suivant:

THÉORÈME 3.4. *Lorsque T est injective, pour qu'un endomorphisme A de H soit un estimateur linéaire admissible de b, il faut et il suffit que A s'écrive sous la forme $A = \bar{A} \circ \mathscr{C}$, où $\mathscr{C}$ est un estimateur des moindres-carrés généralisés de b et $\bar{A}$ un estimateur linéaire admissible de l'application identique de H relativement à la structure statistique induite sur H par $\mathscr{C}$. Lorsque L est inclus dans l'image de V, il équivaut à dire que $\bar{A}$ est un endomorphisme de H vérifiant la condition*

$$\bar{A} \circ ({}^tT \circ V^- \circ T)^{-1} \circ {}^t\bar{A} \preccurlyeq_L \bar{A} \circ ({}^tT \circ V^- \circ T)^{-1},$$

où V^- désigne une g-inverse arbitraire de V.

Remarques. 1. On doit à C. R. Rao la propriété remarquable que, lorsque E est un espace $\boldsymbol{R}^n$, tout estimateur linéaire admissible A de b est l'image linéaire $\bar{A} \circ \mathscr{C}$ d'un estimateur des moindres-carrés généralisés $\mathscr{C}$ de b (cf. théorèmes 6.6 de [17] et 3.4 de [13]). Baksalary et Markiewicz ont également prouvé ce résultat par une autre voie (cf. [4], corollaire 4, pp. 357–358). Nous donnons une caractérisation statistique de l'application linéaire $\bar{A}$ liée au choix de $\mathscr{C}$: $\bar{A}$ est un estimateur linéaire admissible de la moyenne de $\mathscr{C}Y$ dont *le lieu géométrique est tout l'espace H.* A noter que c'est historiquement sous une telle hypothèse que Rao a formulé dans [17] les premières propriétés de l'admissibilité en estimation linéaire.

2. La dernière assertion de ce théorème est en fait vraie chaque fois que le tenseur de variance $\mathscr{V}$ de $\mathscr{C}$ est bijectif (ce qui est réalisé si V est une g-inverse de l'inverse généralisée W^- de W associée à $\mathscr{C}$), à condition de remplacer alors $({}^tT \circ V^- \circ T)^{-1}$ par $\mathscr{V}$.

3.3. L'admissibilité sur un modèle à contraintes structurelles. Supposons maintenant que le paramètre b soit assujetti à décrire un sous-espace affine donné D de H de direction Δ. Nous noterons $[b \in D]$ une telle contrainte que nous qualifierons de *contrainte structurelle*. Soit R une application linéaire de H dans F identifiable sous cette contrainte, autrement dit, telle que, pour tout couple (b_1, b_2) d'éléments de H vérifiant $Tb_1 = Tb_2$, on ait $Rb_1 = Rb_2$. Par une adaptation immédiate du théorème de factorisation des applications linéaires, il équivaut à écrire (cf. [1]), que

$$\Delta \cap \operatorname{Ker} T \subset \operatorname{Ker} R,$$

et donc à dire que R admet un estimateur affine sans biais sous la contrainte $[b \in D]$. Si Ψ désigne une application linéaire d'un espace vectoriel réel $\mathscr{H}$ (que nous supposerons ici euclidien) dans H d'image Δ, l'identifiabilité de R sous la contrainte $[b \in D]$ équivaut à l'identifiabilité de $R \circ \Psi$ relativement au modèle de Gauss–Markov $\mathscr{P}(R \circ \Psi \mathfrak{b}, \{V\})$, $\mathfrak{b}$ désignant le nouveau paramètre de régression dans $\mathscr{H}$.

La recherche d'estimateurs linéaires (ou affines) de R se ramène à la recherche d'estimateurs linéaires admissibles de $R \circ \Psi$ relativement au modèle $\mathscr{P}(R \circ \Psi \mathfrak{b}, \{V\})$: pour qu'un estimateur affine A' de R, de partie linéaire A, soit admissible, il faut et il suffit que A soit un estimateur linéaire admissible de R relativement au modèle $\mathscr{P}(R \circ \Psi \mathfrak{b}, \{V\})$.

Le théorème 2.12 et le corollaire 2.13 conduisent à la proposition suivante qui constitue une version descriptive plus précise du théorème 3 de [2] (p. 15):

PROPOSITION 3.5. *Soient R une application linéaire identifiable sous la contrainte affine $[b \in D]$, et Ψ une application linéaire d'image Δ, direction de D. Alors, pour qu'un estimateur affine $A' = t_a \circ A$ de R soit admissible sous la contrainte $[b \in D]$, il faut et il suffit que les deux conditions suivantes soient satisfaites*:

(i) *A est un estimateur linéaire admissible de R relativement au modèle de Gauss–Markov $\mathscr{P}(R \circ \Psi \mathfrak{b}, \{V\})$*;

(ii) *a appartient au sous-espace affine $(R - A \circ T)(D)$ de F.*

La vérification de l'assertion (ii) est immédiate. Au moyen des résultats établis antérieurement, il est possible de dégager aussitôt une série de propriétés caractéristiques de la famille des estimateurs admissibles A de R relativement au modèle $\mathscr{P}(T \circ \Psi \mathfrak{b}, \{V\})$. Evidemment, R n'admettra des estimateurs linéaires admissibles sous la contrainte $[b \in D]$ que si D est un sous-espace vectoriel de H (cf. corollaire 2.13).

Un cas particulier important concerne encore la recherche d'estimateurs linéaires ou affines admissibles de b sous la contrainte $[b \in D]$ lorsque la restriction T' de T au sous-espace Δ est injective. D'après la proposition précédente, il équivaut à supposer que D est un sous-espace vectoriel Δ de H. Observons que, si $T(\Delta)+\operatorname{Im} V = E$, alors l'image de tout estimateur linéaire admissible A de b sous la contrainte $[b\in\Delta]$ est contenue dans Δ (cf. théorème 2.6) et $T\circ A$ est un estimateur linéaire admissible de μ relativement au modèle $\mathscr{P}[T(\Delta),\{V\}]$ (stabilité par linéarité de l'admissibilité). Usant du même principe, la réciproque est vraie en composant (à gauche) $T\circ A$ par une inverse à gauche de la restriction de T à Δ, si bien que l'on obtient aussitôt la transposition suivante du théorème 3.1:

THÉORÈME 3.6. *Supposons b identifiable sous la contrainte linéaire $[b\in\Delta]$. Alors pour qu'un estimateur linéaire A de b soit admissible sous cette contrainte, il suffit que l'image de A soit contenue dans Δ et que $T\circ A$ soit un estimateur linéaire admissible de μ relativement au modèle $\mathscr{P}[T(\Delta),\{V\}]$. Lorsque $T(\Delta)+\operatorname{Im} V = E$, ces deux conditions sont nécessaires et suffisantes pour que A soit un estimateur linéaire de b sous la contrainte $[b\in\Delta]$.*

TRAVAUX CITÉS

[1] I. S. Alalouf and G. P. H. Styan, *Characterizations of estimability in the general linear models*, Ann. Statist. 7 (1979), pp. 194–200.

[2] J. K. Baksalary and A. Markiewicz, *Admissible linear estimators in restricted linear models*, Linear Algebra Appl. 70 (1985), pp. 9–19.

[3] — *Characterizations of admissible linear estimators in restricted linear models*, J. Statist. Plann. Inference 13 (1986), pp. 395–398.

[4] — *Admissible linear estimators in the general Gauss–Markov model*, ibidem 19 (1988), pp. 349–359.

[5] — *A matrix inequality and admissibility of linear estimators with respect to mean square error matrix criterion*, Linear Algebra Appl. 112 (1989), pp. 9–18.

[6] — *Admissible linear estimators of an arbitrary vector of parametric functions in the general Gauss–Markov model*, J. Statist. Plann. Inference 26 (1990), pp. 161–171.

[7] J. K. Baksalary and Th. Mathew, *Admissible linear estimation in the general Gauss–Markov model with an incorrectly specified dispersion matrix*, J. Multivariate Anal. 27 (1988), pp. 53–67.

[8] W. Klonecki. *Linear estimators of the mean vector in linear models: problem of admissibility*, Probab. Math. Statist. 2 (1982), pp. 167–178.

[9] — and S. Zontek, *On the structure of admissible linear estimators*, J. Multivariate Anal. 24 (1988), pp. 11–30.

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

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

[12] F. Y. Lin and T. S. Yong, *Necessary and sufficient conditions that linear estimators of a mixed effects linear model are admissible under matrix loss function*, Statistics 24 (1993), pp. 303–309.

[13] Th. Mathew, C. R. Rao and B. K. Sinha, *Admissible linear estimation in singular linear models*, Comm. Statist. A — Theory Methods 13 (1984), pp. 3033–3045.

[14] A. Olsen, J. Seely and D. Birkes, *Invariant quadratic unbiased estimation for two variance components*, Ann. Statist. 4 (1976), pp. 878–890.

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

[16] C. R. Rao, *Unified theory of least squares*, Comm. Statist. 1 (1972), pp. 1–8.

[17] — *Estimation of parameters in a linear model*, Ann. Statist. 4 (1976), pp. 1023–1037. *Correction*, ibidem 7 (1979), p. 696.

[18] J.-J. Téchené, *Une introduction à la théorie générale de l'approximation quadratique d'une application linéaire*, Probab. Math. Statist. 15 (1995), pp. 469–492.

[19] — *Sur la notion de projection orthogonale dans un espace semi-euclidien*, Linear Algebra Appl., Fifth Special Issue on Linear Algebra and Statistics (1996), pp. 239–268.

[20] — *Les aspects fondamentaux de l'admissibilité en approximation quadratique d'applications linéaires*, Linear Algebra Appl., Sixth Special Issue on Linear Algebra and Statistics (1997), pp. 389–419.

[21] — *Logique des moindres-carrés et inférence statistique*, Thèse de Doctorat d'Etat, Université de Pau et des Pays de l'Adour, France, 1994.

Département de Mathématiques — Recherche
I.P.R.A. — Université de Pau et des Pays de l'Adour
Avenue de l'Université — 64000 Pau, France

Received on 1.9.1997