

3. Cas ou $\text{rg } A = \text{dim } B$.

Objektyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **6 (1960)**

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **21.07.2024**

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

LES MODÈLES LINÉAIRES EN ANALYSE STATISTIQUE

par H. BRENY

(suite)

3. CAS OU $\text{rg } \mathfrak{A} = \dim \mathbf{B}$.

3, 1. *Remarques théoriques.*

3, 11. Lorsque $r = p$, $\mathfrak{G} \equiv \mathfrak{A}^T \mathfrak{A}$ est une matrice symétrique $p \times p$, régulière, dont on peut former l'inverse \mathfrak{G}^{-1} . Alors, de

$$\mathfrak{G} \hat{\mathbf{b}}_H = \mathfrak{A}^T \mathfrak{x}$$

on tire

$$\hat{\mathbf{b}}_H = \mathfrak{G}^{-1} \mathfrak{A}^T \mathfrak{x},$$

et on en déduit

$$\begin{aligned} \mathbf{C} \hat{\mathbf{b}}_H &= \mathbf{E} (\hat{\mathbf{b}}_H - \mathfrak{b}_H) (\hat{\mathbf{b}}_H - \mathfrak{b}_H)^T \\ &= \mathbf{E} \mathfrak{G}^{-1} \mathfrak{A}^T (\mathfrak{x} - \mathbf{E} \mathfrak{x}) (\mathfrak{x} - \mathbf{E} \mathfrak{x})^T \mathfrak{A} \mathfrak{G}^{-1}. \end{aligned}$$

Mais

$$\mathbf{E} (\mathfrak{x} - \mathbf{E} \mathfrak{x}) (\mathfrak{x} - \mathbf{E} \mathfrak{x})^T = \mathbf{C} \mathfrak{x} = \mathfrak{S}_n \sigma^2,$$

donc

$$\mathbf{C} \hat{\mathbf{b}}_H = \mathfrak{G}^{-1} \mathfrak{A}^T (\mathfrak{S}_n \sigma^2) \mathfrak{A} \mathfrak{G}^{-1} = (\mathfrak{G}^{-1} \mathfrak{A}^T \mathfrak{A} \mathfrak{G}^{-1}) \sigma^2 = \mathfrak{G}^{-1} \sigma^2. \quad (13)$$

Le calcul de \mathfrak{G}^{-1} équivaut donc à celui de $\mathbf{C} \hat{\mathbf{b}}_H$.

3, 12. Pour le calcul de **SCN**, il faut former les équations (7); ici, les u_i^* sont représentés par les lignes de \mathfrak{A}^T , c_i^T , de sorte que, aux inconnues près, le système (7) n'est autre que le système normal (donc, $\lambda_i = \hat{\mathbf{b}}_{H,i}$); il résulte alors de (8) que

$$\begin{aligned} \text{SCN} &= \sum_1^p \hat{\mathbf{b}}_{H,i} c_i^T \mathfrak{x} = \hat{\mathbf{b}}_H^T \mathfrak{A}^T \mathfrak{x} \\ &= (\mathfrak{x}^* \mathfrak{A}) \mathfrak{G}^{-1} (\mathfrak{A}^T \mathfrak{x}). \quad (14) \end{aligned}$$

3, 13. SCE acquiert ici une signification très simple. En effet, si on pose

$$\mathbf{x} = \mathfrak{A} \hat{\mathbf{b}}_H + \mathbf{r}$$

(de sorte que les composantes $r_{H,i}$ de \mathbf{r}_H sont les « résidus »), on a, d'une part,

$$\mathfrak{A}^T \mathbf{x} = \mathfrak{A}^T \mathfrak{A} \hat{\mathbf{b}}_H,$$

et, d'autre part,

$$\mathfrak{A}^T \mathbf{x} = \mathfrak{A}^T \mathfrak{A} \hat{\mathbf{b}}_H + \mathfrak{A}^T \mathbf{r};$$

on a donc $\mathfrak{A}^T \mathbf{r} = 0$; dès lors

$$\begin{aligned} \mathbf{x}^* \mathbf{x} &= (\hat{\mathbf{b}}_H^T \mathfrak{A}^T + \mathbf{r}^T) (\mathfrak{A} \hat{\mathbf{b}}_H + \mathbf{r}) \\ &= \hat{\mathbf{b}}_H^T \mathfrak{A}^T \mathfrak{A} \hat{\mathbf{b}}_H + \mathbf{r}^T \mathbf{r} + \hat{\mathbf{b}}_H^T \mathfrak{A}^T \mathbf{r} + \mathbf{r}^T \mathfrak{A} \hat{\mathbf{b}}_H \\ &= \hat{\mathbf{b}}_H^T \mathfrak{A}^T \mathbf{x} + \mathbf{r}^T \mathbf{r} = \text{SCN} + \mathbf{r}^T \mathbf{r}, \end{aligned}$$

d'où, puisque $\mathbf{x}^* \mathbf{x} = \text{SCT}$,

$$\text{SCE} = \mathbf{r}^T \mathbf{r} = \sum_1^n r_{H,i}^2.$$

3, 14. Si \mathfrak{G} est diagonale (ce qui arrive si et seulement si les lignes de \mathfrak{A}^T correspondent à des vecteurs deux à deux orthogonaux de \mathbf{V}_+), les $\hat{\mathbf{b}}_{H,i}$ sont deux à deux orthogonaux, et

$$\text{SCN} = \sum_1^p \text{SC}\{\hat{\mathbf{b}}_{H,i}\}.$$

3, 2. Exécution des calculs.

3, 21. La résolution des équations normales peut évidemment se faire par un procédé quelconque. On sait toutefois, depuis Benoît et Banachiewicz, que les procédés « compacts » habituels constituent tous des variations plus ou moins heureuses de la méthode de « factorisation triangulaire », particulièrement simple à appliquer dans le cas d'une matrice symétrique, comme l'est \mathfrak{G} (cfr. [VI, VII]).

3, 22. Il s'agit, en principe, de trouver une matrice \mathfrak{S} , triangulaire supérieure, telle que

$$\mathfrak{G} = \mathfrak{S}^T \mathfrak{S} = \mathfrak{S} \times^0 \mathfrak{S}.$$

Le calcul de \mathfrak{S} se fait ligne par ligne, suivant le schéma ¹³⁾

$$\begin{cases} (\mathfrak{S} \varepsilon_i) \times^0 (\mathfrak{S} \varepsilon_j) = \varepsilon_i^T \mathfrak{G} \varepsilon_j, \\ i = 1, j = 1, 2, \dots; i = 2, j = 2, 3, \dots; \text{etc.} \end{cases}$$

Si on désigne par g_Σ la colonne des sommes des éléments des lignes de \mathfrak{G} , et par \mathfrak{s}_Σ la colonne des sommes des éléments des lignes de \mathfrak{S} , on a

$$(\mathfrak{S} \varepsilon_i) \times^0 \sum_{j=1}^p (\mathfrak{S} \varepsilon_j) = \sum_{j=1}^p \varepsilon_i^T \mathfrak{G} \varepsilon_j$$

ou
$$(\mathfrak{S} \varepsilon_i) \times^0 \mathfrak{s}_\Sigma = \varepsilon_i^T g_\Sigma ;$$

chaque ligne de \mathfrak{S} est ainsi vérifiée dès qu'elle est calculée.

3, 23. \mathfrak{S} étant calculée, on a le choix, pour poursuivre, entre deux méthodes; la première fournit SCN et \mathfrak{G}^{-1} (donc, à un facteur près, $\mathbf{C} \hat{\mathbf{b}}_H$); la seconde fournit une décomposition de SCN en termes de la forme red $[\mathbf{U}_1^* \mid \mathbf{U}_2^*, \dots, \mathbf{U}_{i-1}^*]$, les \mathbf{U}_i^* étant orthogonaux et unidimensionnels.

3, 231. Dans la première méthode, on calcule \mathfrak{G}^{-1} sans calculer \mathfrak{S}^{-1} , à partir de la relation $\mathfrak{G}^{-1} = \mathfrak{S}^{-1} (\mathfrak{S}^T)^{-1}$, qui donne

$$\mathfrak{S} \times_0 (\mathfrak{G}^{-1})^T = (\mathfrak{S}^T)^{-1} ;$$

mais, d'une part, $(\mathfrak{G}^{-1})^T = \mathfrak{G}^{-1}$ et, d'autre part, $(\mathfrak{S}^T)^{-1}$ est une matrice triangulaire inférieure, de sorte que

$$\varepsilon_i^T (\mathfrak{S}^T)^{-1} \varepsilon_j = 0 \quad (i < j) , \quad \varepsilon_i^T (\mathfrak{S}^T)^{-1} \varepsilon_i = \varepsilon_i^T \mathfrak{S}^{-1} \varepsilon_i = [\varepsilon_i^T \mathfrak{S} \varepsilon_i]^{-1} .$$

On part donc de

$$(\varepsilon_i^T \mathfrak{S}) \times_0 (\varepsilon_j^T \mathfrak{G}^{-1}) = \varepsilon_i^T (\mathfrak{S}^T)^{-1} \varepsilon_j$$

et on prend successivement

$$\begin{aligned} j = p, \quad i = p, \quad p - 1, \quad p - 2, \quad \dots; \\ j = p - 1, \quad i = p - 1, \quad p - 2, \quad \dots; \\ \text{etc.} \end{aligned}$$

On calcule ensuite aisément

$$\hat{\mathbf{b}}_H^T = (\mathfrak{G}^{-1} \mathfrak{A}^T \mathfrak{r})^T = \mathfrak{G}^{-1} \times_0 (\mathfrak{A}^T \mathfrak{r})^T$$

puis

$$SCN = \hat{b}_H^T \mathfrak{A}^T \mathfrak{r} = \hat{b}_H^T \times_0 (\mathfrak{A}^T \mathfrak{r})^T .$$

Une vérification des calculs est possible: comme $\mathfrak{G}^{-1} \mathfrak{G} = \mathfrak{S}_p$, on a

$$\mathfrak{G}^{-1} \left(\sum_1^p \mathfrak{G} \varepsilon_j \right) = \sum_1^p \mathfrak{S}_p \varepsilon_j = \| 1, \dots, 1 \| ,$$

donc

$$\mathfrak{G}^{-1} \times^0 \mathfrak{g}_\Sigma = \| 1, \dots, 1 \|^T ,$$

ce qui vérifie chaque colonne de \mathfrak{G}^{-1} dès qu'elle est calculée.

3, 2324. Dans la seconde méthode, on pose (cfr. [VIII]), rapportant \mathbf{B} à la base \mathfrak{R} au lieu de \mathfrak{S} :

$$b_K = \mathfrak{S} b_H, \quad \hat{b}_K = \mathfrak{S} \hat{b}_H,$$

d'où

$$\mathfrak{S}^T \hat{b}_K = \mathfrak{S}^T \mathfrak{S} \hat{b}_H = \mathfrak{G} \hat{b}_H = \mathfrak{A}^T \mathfrak{r}$$

ou encore

$$\mathfrak{S} \times^0 \hat{b}_K = \mathfrak{A}^T \mathfrak{r},$$

formule qui se prête immédiatement au calcul numérique.

\hat{b}_K est évidemment l'estimateur privilégié de b_K . Il jouit de la propriété suivante:

$$\begin{aligned} \mathbf{C} b_K &= \mathbf{E} [\mathfrak{S} (\hat{b}_H - b_H)] [\mathfrak{S} (\hat{b}_H - b_H)]^T = \mathfrak{S} [\mathbf{C} \hat{b}_H] \mathfrak{S}^T \\ &= \mathfrak{S} (\mathfrak{G}^{-1} \sigma^2) \mathfrak{S}^T = \mathfrak{S}_p \sigma^2 ; \end{aligned}$$

\hat{b}_K est donc orthonormal, et l'on a

$$\text{var} (\varepsilon_i^T \hat{b}_K) = \sigma^2 ,$$

$$\text{cov} (\varepsilon_i^T \hat{b}_K, \varepsilon_j^T \hat{b}_K) = 0 \quad (i \neq j) ,$$

$$SCN = \sum_1^p \mathbf{SC} \{ \varepsilon_i^T \hat{b}_K \} .$$

Mais, si l'on pose $\varepsilon_i^T \hat{b}_K = l_i^* \mathfrak{r}$, on a $\text{var} (\varepsilon_i^T \hat{b}_K) = \sigma^2 l_i^* l_i$, ce qui montre que $l_i^* l_i = 1$; dès lors, puisque

$$\mathbf{SC} \{ \varepsilon_i^T \hat{b}_K \} = \mathbf{SC} \{ l_i^* \} = (l_i^* \mathfrak{r})^2 / (l_i^* l_i) = (l_i^* \mathfrak{r})^2 ,$$

on a

$$\text{SCN} = \sum_1^p (\varepsilon_i^T \hat{\mathbf{b}}_K)^2. \quad (15)$$

3, 2322. $\hat{\mathbf{b}}_H$ s'obtient immédiatement à partir de $\hat{\mathbf{b}}_K$ par la formule

$$\mathfrak{S} \times_0 \hat{\mathbf{b}}_H^T = \hat{\mathbf{b}}_K,$$

où l'on utilise successivement les lignes de \mathfrak{S} en remontant de la dernière à la première.

Ici aussi, des vérifications sont possibles. On doit en effet avoir

$$(\varepsilon_i^T \mathfrak{S}^T) \hat{\mathbf{b}}_K = \varepsilon_i^T \mathfrak{A}^T \mathfrak{r}$$

d'où

$$\mathfrak{s}_\Sigma \times_0 \hat{\mathbf{b}}_K = \sum_i \varepsilon_i^T \mathfrak{A}^T \mathfrak{r},$$

ce qui vérifie le calcul de $\hat{\mathbf{b}}_K$. Si, ensuite, on appelle \mathfrak{G}_Σ^T la ligne des sommes des termes des colonnes de \mathfrak{S} , on doit avoir

$$\mathfrak{G}_\Sigma^T \times_0 \hat{\mathbf{b}}_H^T = \sum_1^p \varepsilon_i^T \hat{\mathbf{b}}_K,$$

ce qui vérifie le calcul de $\hat{\mathbf{b}}_H^T$.

3, 2323. Considérons la matrice \mathfrak{S}_t formée par les t premières lignes et colonnes de \mathfrak{S} ; le mode de calcul de \mathfrak{S} montre immédiatement que \mathfrak{S}_t ne dépend que de la matrice \mathfrak{G}_t formée des t premières lignes et colonnes de \mathfrak{G} . Par contre, $(\mathfrak{G}_t)^{-1}$ ne dépend pas seulement de \mathfrak{S}_t (le calcul de \mathfrak{G}^{-1} commence par le coin inférieur droit). De même, $\varepsilon_1^T \hat{\mathbf{b}}_K, \dots, \varepsilon_t^T \hat{\mathbf{b}}_K$ ne dépendent que de \mathfrak{S}_t , mais $\varepsilon_1^T \hat{\mathbf{b}}_H, \dots, \varepsilon_t^T \hat{\mathbf{b}}_H$ dépendent de l'ensemble de \mathfrak{S} (le calcul de $\hat{\mathbf{b}}_H$ commence par $\varepsilon_p^T \hat{\mathbf{b}}_H$). Dès lors, supposons que le modèle soit réduit à ses t premiers paramètres, $\varepsilon_1^T \hat{\mathbf{b}}_H, \dots, \varepsilon_t^T \hat{\mathbf{b}}_H$, ce qui revient à remplacer \mathfrak{G} par \mathfrak{G}_t et \mathbf{b}_H par $[\varepsilon_1^T \hat{\mathbf{b}}_H, \dots, \varepsilon_t^T \hat{\mathbf{b}}_H]^T$, ou encore à supposer $\varepsilon_{t+1}^T \hat{\mathbf{b}}_H = \dots = \varepsilon_p^T \hat{\mathbf{b}}_H = 0$; ce modèle est donc caractérisé par le fait que

$$(l^* \in \{\varepsilon_{t+1}^*, \dots, \varepsilon_p^*\}) \quad \text{implique} \quad \mathbf{E} l^* \mathfrak{r} = 0, \quad (16)$$

relation de la forme (11). Dans ce modèle réduit, rien n'est changé au calcul de $\varepsilon_1^T \hat{\mathbf{b}}_K, \dots, \varepsilon_t^T \hat{\mathbf{b}}_K$ et, si l'on note SCN_t et SCE_{n-t} res-

pectivement la somme de carrés normale et la somme de carrés de l'erreur de ce modèle, on a, en vertu de (15),

$$\mathbf{SC}N_t = \mathbf{SC}\{e_1^*, \dots, e_t^*\} = \sum_1^t (\varepsilon_i^T \hat{\mathbf{b}}_K)^2,$$

donc

$$\mathbf{SCE}_{n-t} = \mathbf{SCT} - \sum_1^t (\varepsilon_i^T \hat{\mathbf{b}}_K)^2,$$

$$\mathbf{SCE}_{n-t-1} = \mathbf{SCT} - \sum_1^{t+1} (\varepsilon_i^T \hat{\mathbf{b}}_K)^2,$$

ce qui implique

$$\mathbf{red} [\{e_{t+1}^*\} | \{e_1^*, \dots, e_t^*\}] = (\varepsilon_{t+1}^T \hat{\mathbf{b}}_K)^2.$$

La notation du premier membre de cette formule est entièrement comparable à celles du § 2, 4; elle est toutefois un peu lourde, et on la remplace le plus souvent par

$$\mathbf{red} [\varepsilon_{t+1}^T \mathbf{b}_H | \varepsilon_1^T \mathbf{b}_H, \dots, \varepsilon_t^T \mathbf{b}_H] \equiv \mathbf{red} [\{e_{t+1}^*\} | \{e_1^*, \dots, e_t^*\}].$$

Dans ces conditions, la table d'analyse de variance s'écrit ainsi:

Sommes de carrés	Formules	Degrés de liberté
\mathbf{SCT}	$\sum_1^n x_i^2$	n
\mathbf{SCN}	$\sum_1^p (\varepsilon_i^T \hat{\mathbf{b}}_K)^2$	p
$\left\{ \begin{array}{l} \mathbf{red} [\varepsilon_1^T \mathbf{b}_H] \\ \mathbf{red} [\varepsilon_2^T \mathbf{b}_H \varepsilon_1^T \mathbf{b}_H] \\ \vdots \\ \mathbf{red} [\varepsilon_p^T \mathbf{b}_H \varepsilon_1^T \mathbf{b}_H, \dots, \varepsilon_{p-1}^T \mathbf{b}_H] \end{array} \right.$	$\left\{ \begin{array}{l} (\varepsilon_1^T \hat{\mathbf{b}}_K)^2 \\ (\varepsilon_2^T \hat{\mathbf{b}}_K)^2 \\ \vdots \\ (\varepsilon_p^T \hat{\mathbf{b}}_K)^2 \end{array} \right.$	$\left\{ \begin{array}{l} 1 \\ 1 \\ \vdots \\ 1 \end{array} \right.$
\mathbf{SCE}	$\mathbf{SCT} - \mathbf{SCN}$	$n - p$
$\left\{ \begin{array}{l} \mathbf{SCint} \\ \mathbf{SCEM} \end{array} \right.$	$\left\{ \begin{array}{l} ? \\ \mathbf{SCE} - \mathbf{SCint} \end{array} \right.$	$\left\{ \begin{array}{l} u \\ n - p - u \end{array} \right.$

3, 24. La disposition matérielle des calculs a une certaine importance; le cas $p = 3$ est décrit en détail ci-dessous.

On calcule et vérifie \mathfrak{S} comme ci-dessus, puis:

$$\begin{aligned}
 x a_1 &= n & y a_1 + u a_2 &= p & z a_1 + v a_2 + w a_3 &= q \\
 A &= a_1 + a_2 + a_3 & T &= n + p + q & t a_1 + r a_2 + s a_3 &= T \\
 c_1 &= x & c_2 &= y + u & c_3 &= z + v + w \\
 w b_3 &= a_3 & v b_3 + u b_2 &= a_2 & z b_3 + y b_2 + x b_1 &= a_1 \\
 b_1 c_1 + b_2 c_2 + b_3 c_3 &= A \\
 SCN &= a_1^2 + a_2^2 + a_3^2 .
 \end{aligned}$$

4. EXEMPLES.

4, 1. Les épreuves de Student.

4, 11. Soit $\mathbf{x}_1, \dots, \mathbf{x}_n$ un échantillon simple et fortuit d'une population normale de moyenne μ et écart-type σ . La théorie des modèles linéaires s'applique ici, avec

$$r = p = 1, \quad \mathfrak{X} = \|\mathbf{1}, \dots, \mathbf{1}\| \quad b_H = \|\mu\|,$$

$$\mathfrak{X}^T \mathfrak{X} = \|n\|, \quad \mathfrak{X}^T \mathfrak{x} = \sum_1^n \mathbf{x}_i,$$

et le système normal se réduit à

$$n \hat{\mu} = \sum_1^n \mathbf{x}_i, \quad \hat{\mu} = \frac{1}{n} \sum_1^n \mathbf{x}_i = \mathbf{m}.$$

On a alors

$$SCN = (\sum \mathbf{x}_i)^2/n, \quad SCT = \sum_1^n \mathbf{x}_i^2,$$

$$SCE = \sum_1^n \mathbf{x}_i^2 - \left(\sum_1^n \mathbf{x}_i \right)^2/n = \sum_1^n (\mathbf{x}_i - \mathbf{m})^2 = \frac{n \sum_1^n \mathbf{x}_i^2 - \left(\sum_1^n \mathbf{x}_i \right)^2}{n}.$$

Si $\mu = a$, l'expression

$$\frac{(\mathbf{m} - a) \sqrt{(n-1)}}{\sqrt{(SCE/n)}} = \sqrt{(n-1)} \frac{\sum_1^n \mathbf{x}_i - na}{\sqrt{(n SCE)}}$$

est une aléatoire t_{n-1} ; SCE/σ^2 est une aléatoire χ_{n-1}^2 .