Statistique Mathématique 2 (MATH-F-309, Chapitre #6)

Thomas Verdebout

Université Libre de Bruxelles

2015/2016

Plan du cours

1. Vecteurs aléatoires.
2. Loi normale multivariée.
3. Inférence dans les modèles gaussiens.
4. Méthodes classiques de l'analyse multivariée.
5. Données directionelles.
6. $\boxed{\text{Modèle linéaire}}$

$\boxed{\text{6. Modèles linéaires.}}$

$\boxed{\text{6.1. Hypothèse de linéarité.}}$

6.2. MLE.

6.3. Une approche par projection.

6.4. Théorie distributionnelle.

6.5. Inférence.

6.6. Analyse des residus.

L'hypothèse linéaire

Définition

Soit Y un n-vecteur aléatoire avec $EY = \mu$. Un modèle linéaire de dimension k est une hypothèse pour μ de la forme $H_0 : \mu \in L$, où L est un sous-espace linéaire de $\mathbb{R}^n$ avec $\mathrm{dim(L) = k}$.

Pour une base donnée $X = (x_1, \ldots, x_k)$ de L, ceci veut donc dire que sous H_0 il existe un vecteur de coefficients $\beta = (\beta_1, \ldots, \beta_k)'$ tel que

$$\mu = \sum_{i=1}^{k} \beta_i x_i = X\beta.$$

Autrement exprimé,

$$Y = X\beta + \varepsilon,$$

où $E\varepsilon = 0$ (et assez souvent on suppose que $\mathrm{Var}(\varepsilon) = \sigma^2 I_n$).

L’hypothèse linéaire

Il est à remarquer que d’après la manière dont ce modèle est défini, on ne l’attribue en général pas au domaine de l’analyse multivariée. La raison en est que $Y = (Y_1, \ldots, Y_n)'$ constitue notre échantillon (univarié).

Il existe aussi une version multivariée du modèle linéaire général. Alors les Y_i' sont des p-vecteurs aléatoires et nous avons une hypothèse de la forme

$$EY = M = XB,$$

où $M \in \mathbb{R}^{n \times p}$, $X \in \mathbb{R}^{n \times q}$ et $B \in \mathbb{R}^{q \times p}$.

L'hypothèse linéaire

Exemples.

- Les Y_i sont i.i.d. de moyenne μ.

  Alors $X = (1, \ldots, 1)'$ et $\beta = \mu$. L'estimation et l'inférence se réduisent à des procédures bien connues.
- La moyenne de Y_i dépend de plusieurs variables (connues) (*variables indépendantes, régresseurs*) $x_1, \ldots, x_k$ de façon linéaire:

$$Y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + \cdots + x_{ki}\beta_k + \varepsilon_i.$$

L'hypothèse linéaire

Principaux objectifs:

- Estimer les paramètres inconnus β (ou μ respectivement) et σ^2 si $\mathrm{Var}(\varepsilon) = \sigma^2 I_n$.
- Etablir la distribution des paramètres quand Y est un vecteur multinormal.
- Tester différentes hypothèses pour β (ou μ respectivement). P.ex. $\beta_\ell = 0$ v.s. $\beta_\ell \neq 0$ ou $\beta_1 = \beta_2 = \cdots = \beta_k$ v.s. ils ne sont pas tous égaux ou $\mu \in L_2$ v.s. $\mu \in L_1 \backslash L_2$ ...
- Verifier si le modèle est approprié.

$\boxed{\text{6. Modèles linéaires.}}$

6.1. Hypothèse de linéarité.

$\boxed{\text{6.2. MLE.}}$

6.3. Une approche par projection.

6.4. Théorie distributionnelle.

6.5. Inférence.

6.6. Analyse des residus.

MLE

Supposons que $Y \sim \mathcal{N}_n(\mu, \sigma^2 I_n)$. Alors la fonction de vraisemblance est donnée par

$$\begin{aligned}\ell(\mu, \sigma | Y) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}(Y-\mu)'(Y-\mu)\right) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\|Y-\mu\|^2\right).\end{aligned}$$

Si le modèle linéaire tient pour $Y \sim \mathcal{N}_n(X\beta, \sigma^2 I_n)$ alors le MLE pour (β, σ^2) équivaut à

$$(\hat{\beta}, \hat{\sigma}^2) = \operatorname{argmax}_{\beta, \sigma^2} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\|Y - X\beta\|^2\right).$$

MLE

Notons que

$$(\hat{\beta}, \hat{\sigma}^2) = \mathrm{argmax}_{\beta,\sigma^2} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\|Y - X\beta\|^2\right)$$
$$= \mathrm{argmin}_{\beta,\sigma^2} \left(\frac{n}{2}\log(\sigma^2) + \frac{1}{2\sigma^2}\|Y - X\beta\|^2\right).$$

Si nous fixons σ^2 pour le moment, alors nous devons trouver

$$\hat{\beta} = \mathrm{argmin}_{\beta}\|Y - X\beta\|^2.$$

En posant $\partial\|Y - X\beta\|^2/\partial\beta = 0$, on obtient facilement (exercice) que

$$\hat{\beta} = (X'X)^{-1}X'Y.$$

MLE

Afin d'obtenir $\hat{\sigma}^2$ nous déterminons

$$\mathrm{argmin}_{s>0}\left(\frac{n\log(s)}{2}+\frac{1}{2s}\|Y-X\hat{\beta}\|^2\right).$$

Des calculs directs (exercice) montrent que

$$\hat{\sigma}^2=\frac{1}{n}\|Y-X\hat{\beta}\|^2.$$

MLE

Théorème

Supposons qu'on a un modèle linéaire de la forme $Y \sim \mathcal{N}_n(X\beta, \sigma^2 I_n)$. *Alors le MLE est donné par*

$$\hat{\beta} = (X'X)^{-1}X'Y \quad et \quad \hat{\sigma}^2 = \frac{1}{n}\|Y - X\hat{\beta}\|^2.$$

Le maximum de la vraisemblance est donné par

$$\left(\frac{1}{2\pi e}\right)^{n/2} \left(\hat{\sigma}^2\right)^{-n/2}.$$

$\square$

MLE

Remarque. Les estimateurs

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{n}\|Y - X\hat{\beta}\|^2$$

ne sont pas sans intérêt quand Y n'est pas normalement distribué. L'estimateur $\hat{\beta}$ est aussi appelé *estimateur des moindres carrés.*

Le vecteur $\hat{\varepsilon} := Y - X\hat{\beta}$ est appelé *vecteur des résidus*. Nous notons

$$SSR = \|Y - X\hat{\beta}\|^2$$

la somme des carrés des résidus.

$\boxed{\text{6. Modèles linéaires.}}$

6.1. Hypothèse de linéarité.

6.2. MLE.

$\boxed{\text{6.3. Une approche par projection.}}$

6.4. Théorie distributionnelle.

6.5. Inférence.

6.6. Analyse des residus.

Une approche par projection

Une autre manière d'obtenir le MLE pour β consiste à écrire

$$\hat{\mu} = X\hat{\beta} = \mathrm{argmin}_{\mu \in L} \| Y - \mu \|,$$

où L est l'espace linéaire engendré par les vecteurs $(x_1, \ldots, x_k)$. Par définition $\hat{\mu} = p_L(Y)$, la *projection* de Y sur L.

Notons que la projection $p_L(Y)$ est caractérisée comme l'unique élément dans L tel que

$$\langle Y - p_L(Y), x \rangle = 0 \quad \text{pour tout} \quad x \in L$$

$\Longleftrightarrow$

$$\langle Y - p_L(Y), x_j \rangle = 0 \quad \text{pour tout} \quad j \in \{1, \ldots, k\}$$

$\Longleftrightarrow$

$$X'(Y - p_L(Y)) = 0.$$

Une approche par projection

Il est limpide que $X(X'X)^{-1}X'Y \in L$ et on voit facilement que

$$X'(Y - X(X'X)^{-1}X'Y) = 0.$$

D'où nous avons que

$$\hat{\mu} = p_L(Y) = X(X'X)^{-1}X'Y = X\hat{\beta}.$$

La matrice

$$H = X(X'X)^{-1}X'$$

est appelée *matrice de projection* ou *matrice chapeau*. Elle est (exercice)

- symétrique: $H = H'$, et
- idempotente: $H^2 = H$.

6. Modèles linéaires.

La proposition suivante tient sans hypothèse de normalité.

Proposition

Supposons que Y est un modèle linéaire de dimension k de la forme

$$Y = \mu + \varepsilon = X\beta + \varepsilon \quad \varepsilon \sim (0, \sigma^2 I_n).$$

Soient $\hat{\mu}$, $\hat{\beta}$ et $\hat{\sigma}^2$ définis comme avant. Alors

$$\begin{aligned} &E\hat{\beta} = \beta \quad \textit{et} \quad E\hat{\mu} = \mu; \\ &\mathrm{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} \quad \textit{et} \quad \mathrm{Var}(\hat{\mu}) = \sigma^2 H; \\ &E\hat{\sigma}^2 = \frac{n-k}{n}\sigma^2. \end{aligned}$$

Théorie distributionnelle

Preuve. La moyenne et la variance de $\hat{\beta}$ et $\hat{\mu}$ sont simples. Nous allons calculer l'espérance de $\hat{\sigma}^2$. En ayant recours à $H' = H$ et $H^2 = H$, nous obtenons que la même chose est vraie pour $I_n - H$. Notez aussi que $H\mu = \mu$ et donc $(I_n - H)\mu = 0$. Par conséquent

$$\begin{aligned} E\hat{\sigma}^2 &= \frac{1}{n} E\|(I_n - H)Y\|^2 = \frac{1}{n} E\|(I_n - H)\varepsilon\|^2 \\ &= \frac{1}{n} E\,\mathrm{tr}(\varepsilon'(I_n - H)'(I_n - H)\varepsilon) \\ &= \frac{1}{n}\mathrm{tr}(E\varepsilon\varepsilon'(I_n - H)) = \frac{\sigma^2}{n}\mathrm{tr}(I_n - H). \end{aligned}$$

La preuve suit alors du fait que $\mathrm{tr}(H) = \mathrm{tr}(X(X'X)^{-1}X') = \mathrm{tr}(X'X(X'X)^{-1}) = k$. $\square$

Proposition

Supposons que Y est un modèle linéaire de dimension k de la forme

$$Y = \mu + \varepsilon = X\beta + \varepsilon \quad \varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n).$$

Soient $\hat{\mu}$, $\hat{\beta}$ et $\hat{\sigma}^2$ définis comme avant. Alors

$$
\begin{array}{rl}
(i) & \hat{\beta} \sim \mathcal{N}_p(\beta, \sigma^2 (X'X)^{-1}); \\
(ii) & \hat{\mu} \sim \mathcal{N}_n(\mu, \sigma^2 H); \\
(iii) & \hat{\sigma}^2 \sim \dfrac{\sigma^2}{n} \chi^2_{n-k}; \\
(iv) & \hat{\sigma}^2 \text{ et } \hat{\beta} \text{ (ou } \hat{\mu}\text{, respectivement) sont indépendants.}
\end{array}
$$

Preuve de (i) et (ii). Les vecteurs $\hat{\beta}$ et $\hat{\mu}$ sont des transformations linéaires du vecteur normal Y. Donc par les résultats du Chapitre 2, $\hat{\beta}$ et $\hat{\mu}$ sont de nouveau normalement distribués. Par la proposition ci-dessus, nous connaissons leur moyenne et variance appropriées. Nous obtenons (i) et (ii). $\square$

Les relations (iii) et (iv) suivront des considérations ci-dessous.

Sous-espaces orthogonaux. Deux sous-espaces linéaires L_1 et L_2 de $\mathbb{R}^n$ sont appelés orthogonaux si $x \perp y$ pour tout $x \in L_1$ et $y \in L_2$. Alors nous définissons

$$L_1 \oplus L_2 = \{x + y | x \in L_1 \text{ et } y \in L_2\}.$$

$$L_1 \ominus L_2 = \{x \in L_1 | x \perp L_2\} \quad \text{(en supposant que } L_2 \subset L_1).$$

Lemme

Supposons que $L_2 \subset L_1$ sont des sous-espaces linéaires de $\mathbb{R}^n$. Alors

$$
\begin{aligned}
(i) &\quad L_1 \ominus L_2 \quad \textit{est une sous-espace linéaire};\\
(ii) &\quad L_1 \ominus L_2 \perp L_2;\\
(iii) &\quad (L_1 \ominus L_2) \oplus L_2 = L_1 \quad (\Longrightarrow L^{\perp} \oplus L = \mathbb{R}^n);\\
(iv) &\quad p_{L^{\perp}}(v) = v - p_L(v).
\end{aligned}
$$

Théorie distributionnelle

Proposition

Soient $L_1, L_2, \ldots, L_r$ *des sous-espaces orthogonaux de* $\mathbb{R}^n$ *avec* $\dim(L_i) = k_i$ *et* $L_1 \oplus L_2 \oplus \cdots \oplus L_r = \mathbb{R}^n$ *et soit* $\varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n)$*. Alors*

$$(a) \quad p_{L_1}(\varepsilon), \ldots, p_{L_r}(\varepsilon) \quad \textit{sont indépendants.}$$

$$(b) \quad \|p_{L_i}(\varepsilon)\|^2 \sim \sigma^2 \chi^2_{k_i} \quad 1 \leq i \leq r.$$

Preuve de (a). Comme $L_1, L_2, \ldots, L_r$ sont des sous-espaces orthogonaux de $\mathbb{R}^n$ avec $\dim(L_i) = k_i$ et $L_1 \oplus L_2 \oplus \cdots \oplus L_r = \mathbb{R}^n$, nous avons une base orthonormale (BON)

$$\{e_{ij},\, 1 \le j \le k_i,\; 1 \le i \le r\}$$

de $\mathbb{R}^n$ tel que $E_i = \{e_{ij},\, 1 \le j \le k_i\}$ est une BON de L_i. Il s'ensuit de $E_i' E_i = I_{k_i}$ que

$$p_{L_i}(\varepsilon) = E_i (E_i' E_i)^{-1} E_i' \varepsilon = E_i E_i' \varepsilon.$$

Donc, comme $E_i' E_j = 0$ pour $i \neq j$, nous obtenons

$$\mathrm{Cov}(p_{L_i}(\varepsilon), p_{L_j}(\varepsilon)) = E_i E_i' \mathrm{Cov}(\varepsilon, \varepsilon) E_j E_j' = 0.$$

Preuve de (b). Nous déduisons par le théorème de Pythagore que

$$\begin{aligned}\|p_{L_i}(\varepsilon)\|^2 &= \|\sum_{j=1}^{k_i}\langle e_{ij},\varepsilon\rangle e_{ij}\|^2\\ &= \sum_{j=1}^{k_i}\langle e_{ij},\varepsilon\rangle^2.\end{aligned}$$

Puisque les e_{ij} sont orthonormaux, il suit que les variables $\langle e_{ij},\varepsilon\rangle$ sont des variables aléatoires indépendantes de loi normale. Leur moyenne est 0 et la variance σ^2. $\square$

Preuve de (iii) et (iv) de la proposition précédente. Rappelons que

$$\hat{\sigma}^2 = \frac{1}{n}\|Y - \hat{\mu}\|^2.$$

Nous rappelons que

$$\|Y - \hat{\mu}\|^2 = \|(I_n - H)\varepsilon\|^2.$$

Par le point (iv) du lemme précédent il découle que $(I_n - H)v = p_{L^\perp}(v)$, la projection de v sur l'espace

$$L^\perp = \{x \in \mathbb{R}^n | x \perp L = \mathrm{span}(x_1, \ldots, x_k)\}.$$

Comme $\dim(L^\perp) = n - k$, la preuve de (iii) suit du point (b) de la proposition précédente.

Théorie distributionnelle

Comme

$$\hat{\mu} = HY = X\beta + H\varepsilon = X\beta + p_L(\varepsilon)$$

et

$$\hat{\sigma}^2 = \frac{1}{n}\|(I_n - H)\varepsilon\|^2 = \frac{1}{n}\|p_{L^\perp}(\varepsilon)\|^2,$$

nous concluons du point (a) de la proposition précédente que $\hat{\mu}$ et $\hat{\sigma}^2$ sont indépendants.

L'indṕendance de $\hat{\beta}$ et $\hat{\sigma}^2$ decoule de

$$\hat{\beta} = (X'X)^{-1}X'\hat{\mu}.$$

$\square$

$\boxed{\text{6. Modèles linéaires.}}$

6.1. Hypothèse de linéarité.

6.2. MLE.

6.3. Une approche par projection.

6.4. Théorie distributionnelle.

$\boxed{\text{6.5. Inférence.}}$

6.6. Analyse des residus.

Inférence

Pour un modèle linéaire $Y \sim \mathcal{N}_n(\mu, \sigma^2 I_n)$, nous allons tester des hypothèses de la forme suivante:

$$H_0 : \mu = \sum_{i=1}^{k} x_i \beta_i,$$

contre

$$H_1 : \mu = \sum_{i=1}^{\ell} x_i \beta_i, \quad k < \ell \leq n.$$

Ou de manière plus générale

$$H_0 : \mu \in L_2 \quad \text{contre} \quad H_1 : \mu \in L_1 \backslash L_2 \quad (L_2 \subset L_1).$$

Nous supposons dans la suite que $\dim(L_1) = k_1$ et $\dim(L_2) = k_2 \leq k_1$.

Inférence

Le rapport de vraisemblance est donné par

$$\Lambda(Y) = \frac{\sup_{\mu_2 \in L_2, \sigma_2 > 0} \ell(\mu, \sigma | Y)}{\sup_{\mu_1 \in L_1, \sigma_1 > 0} \ell(\mu, \sigma | Y)}.$$

Rappelons que le maximum de la vraisemblance pour le modèle L est proportionnel à

$$\|Y - p_L(Y)\|^{-n};$$

la statistique du test de rapport de vraisemblance est donnée par

$$\Lambda(Y) = \left(\frac{\|Y - p_{L_1}(Y)\|^2}{\|Y - p_{L_2}(Y)\|^2} \right)^{n/2},$$

et rejette l'hypothèse nulle si $\Lambda(Y)$ est trop petit.

Inférence

Maintenant comme

$$\|Y - p_{L_2}(Y)\|^2 = \|Y - p_{L_1}(Y) + \underbrace{p_{L_1}(Y) - p_{L_2}(Y)}_{\in L_1}\|^2$$

nous obtenons par le théorème de projection et le théorème de Pythagore que

$$\|Y - p_{L_2}(Y)\|^2 = \|Y - p_{L_1}(Y)\|^2 + \|p_{L_1}(Y) - p_{L_2}(Y)\|^2.$$

Donc le test de rapport de vraisemblance est équivalent au test qui rejette l'hypothèse nulle si

$$F(Y) = \frac{\|p_{L_1}(Y) - p_{L_2}(Y)\|^2/(k_1 - k_2)}{\|Y - p_{L_1}(Y)\|^2/(n - k_1)}$$

est trop grand.

Inférence

Maintenant sous H_0 : $\mu \in L_2$ nous avons

$$
\begin{aligned}
F(Y) &= \frac{\|p_{L_1}(Y) - p_{L_2}(Y)\|^2/(k_1 - k_2)}{\|Y - p_{L_1}(Y)\|^2/(n - k_1)} \\
&= \frac{\|p_{L_1}(\varepsilon) - p_{L_2}(\varepsilon)/(k_1 - k_2)\|^2}{\|\varepsilon - p_{L_1}(\varepsilon)\|^2/(n - k_1)} \\
&= \frac{\|p_{L_1 \ominus L_2}(\varepsilon)\|^2/(k_1 - k_2)}{\|p_{L_1^\perp}(\varepsilon)\|^2/(n - k_1)}.
\end{aligned}
$$

Comme $\mathbb{R}^n = L_1^\perp \oplus (L_1 \ominus L_2) \oplus L_2$ nous concluons que

$$
\|p_{L_1 \ominus L_2}(\varepsilon)\|^2 \sim \chi^2_{k_1 - k_2} \quad \perp\!\!\!\perp \quad \|p_{L_1^\perp}(\varepsilon)\|^2 \sim \chi^2_{n - k_1}.
$$

Théorème

Supposons qu'on a un modèle linéaire de la forme $Y \sim \mathcal{N}_n(\mu, \sigma^2 I_n)$. *Le test de rapport de vraisemblance pour* $H_0 : \mu \in L_2$ *contre* $H_1 : \mu \in L_1$ *est équivalent à un test* F. *La statistique de test*

$$F(Y) = \frac{\|p_{L_1}(Y) - p_{L_2}(Y)\|^2/(k_1 - k_2)}{\|Y - p_{L_1}(Y)\|^2/(n - k_1)}$$

suit une distribution F *de* $k_1 - k_2$ *et* $n - k_1$ *degrés de liberté. Pour un test de niveau* α, *nous rejetons* H_0 *si*

$$F(Y) \geq F_{k_1 - k_2, n - k_1; 1-\alpha}.$$

Inférence

Remarque. Définissons
$$\tilde{\sigma}_1^2 = \frac{1}{n-k_1} \|Y - p_{L_1}(Y)\|^2.$$
Remarquons que ceci est un estimateur sans biais pour σ^2 pour le modèle L_1, et que sous H_0 aussi bien que sous H_1 il converge vers σ^2.

Alors nous pouvons écrire
$$F(Y) = \frac{1}{k_1 - k_2} \frac{\|\hat{\mu}_1 - \hat{\mu}_2\|^2}{\tilde{\sigma}_1^2},$$
où les $\hat{\mu}_i$ sont les MLE sous le modèle L_i.

Asymptotiquement nous voyons que $\|\hat{\mu}_1 - \hat{\mu}_2\|^2 \sim \sigma^2 \chi^2_{k_1 - k_2}$.

6. Modèles linéaires.

6.1. Hypothèse de linéarité.

6.2. MLE.

6.3. Une approche par projection.

6.4. Théorie distributionnelle.

6.5. Inférence.

6.6. Analyse des residus.

Analyse des résidus

L'idée générale de l'analyse des résidus réside dans l'étude de divers plots de résidus afin d'évaluer à quel point un modèle donné est proche des/fitte les données. Nous sommes donc amenés à d'abord étudier la distribution des résidus afin de savoir comment ils sont censés se comporter.

Supposons avoir un modèle linéaire $Y \sim \mathcal{N}_n(\mu, \sigma^2 I_n)$, avec $\mu \in L$ sous H_0. Le vecteur de résidus est

$$R = Y - \hat{\mu} = Y - HY = (I_n - H)Y = (I_n - H)(\mu + \varepsilon) \overset{H_0}{=} (I_n - H)\varepsilon.$$

Donc

$$ER = 0 \quad \text{et} \quad \mathrm{Var}(R) = \sigma^2(I_n - H),$$

montrant que

$$R \sim \mathcal{N}_n(0, \sigma^2(I_n - H)).$$

Remarquons que $\hat{\mu}$ et R sont indépendants.

Analyse des résidus

Remarquons que R est dans un certain sens un estimateur pour l'erreur statistique non-observée ε. Comme

$$\varepsilon - R = H\varepsilon,$$

nous avons (nous utilisons encore une fois le fait que H est idempotent) que

$$\mathrm{Var}(\varepsilon - R) = \sigma^2 H.$$

Exemple. Considérons le modèle de régression linéaire simple

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2),$$

où $1 \leq i \leq n$. Nous pouvons réécrire ceci comme

$$Y_i = \alpha_0 + \beta_1 t_i + \varepsilon_i,$$

où $\alpha_0 = \beta_0 + \beta_1 \bar{x}$ et $t_i = x_i - \bar{x}$. Alors $\bar{t} = 0$. Notons $t_+ = \sum t_i^2$.

En forme vectorielle notre modèle correspond à

$$Y = X\beta + \varepsilon, \quad \text{avec} \quad X = (1_n, t) \quad \text{et} \quad \beta = (\alpha_0, \beta_1),$$

où $1_n = (1, \ldots, 1)' \in \mathbb{R}^n$. Notons que $1_n' t = n\bar{t} = 0$.

Exemple, suite.

Dans cet exemple $X'X = \begin{pmatrix} n & 0 \\ 0 & t_+ \end{pmatrix}$. Ceci montre (pourquoi?) que $\hat{\alpha}_0$ et $\hat{\beta}_1$ sont indépendants. De plus nous avons

$$H = X(X'X)^{-1}X' = \left(\left(\frac{1}{n} + \frac{t_i t_j}{t_+}\right)\right)_{i,j=1}^{n}.$$

Si nous supposons que les x_i sont des nombres aléatoires iid, de moyenne m_x et de variance σ_x^2, alors par la loi des grands nombres $H_{ij} = O_P\left(n^{-1}\right)$ pour tout $1 \leq i \leq j \leq n$.

En particulier les éléments diagonaux de H, $\varepsilon_i - R_i$, seront très proches de zéro si n est grand, montrant que les résidus devraient se comporter similairement que ε.

Analyse des résidus

Exemple, suite.

De plus, comme

$$\mathrm{Var}(R) = \sigma^2(I_n - H),$$

nous voyons que

$$\mathrm{cov}(R_i, R_j) \approx \begin{cases} \sigma^2 & \text{si } i = j \\ 0 & \text{autrement} \end{cases}$$

quand n est grand. Ceci est en accord avec $\mathrm{Var}(\varepsilon) = \sigma^2 I_n$.