MAT7381 Exercices
Chapitre 3 - Modèle linéaire général

Notez bien : À moins d'avis contraire la matrice de covariance du vecteur d'observations y est σ^2I.

3.1 Les données suivantes présentent les pertes en humidité dans le sol des forêts exploitées. Les trois traitements diffèrent selon la quantité de bois laissée sur le sol après la coupe. *Traitement* 1: pas de bois; *Traitement* 2: quantité modérée de bois; *Traitement* 3: grande quantité de bois.

Traitement		
1	2	3
1,52	1,63	2,56
1,38	1,82	3,32
	1,35	2,76
	1,03	

Le modèle d'analyse de variance classique peut s'écrire comme

$$y_{ij} = \mu_i + \varepsilon_{ij},\ j = 1, \ldots, n_i \text{ et } i = 1,2,3$$

où y_{ij} est la j^e observation du groupe i. Soit $\boldsymbol{y}$ = [1,52 ; 1,38 ; 1,63 ; 1,82 ; 1,35 ; 1,03 ; 2,56 ; 3,32 ; 2,76]', un vecteur aléatoire de moyenne $[\mu_1 ; \mu_2 ; \mu_2 ;. \mu_2 ; \mu_2 ; \mu_3 ; \mu_3 ; \mu_3]'$. Soit et $\boldsymbol{\mu}' = [\mu_1 ; \mu_2 ; \mu_3]$ et H_o l'hypothèse $H_o : \mu_1 = \mu_2 = \mu_3$.

a) Exprimer l'hypothèse H_o sous sa forme matricielle, $H_o = \mathbf{L}'\boldsymbol{\mu}$.

b) Le modèle peut alors s'écrire comme $\boldsymbol{y} = \mathbf{X}\boldsymbol{\mu} + \boldsymbol{\varepsilon}$. Identifiez la matrice $\mathbf{X}$.

c) Une approche traditionnelle consiste à définir un vecteur $\boldsymbol{\delta}' = [\mu ; \alpha_1 ; \alpha_2 ; \alpha_3]$ de quatre paramètres, où μ est une moyenne globale et α_i est l'écart entre la moyenne du i^e traitement et la moyenne globale. L'expression du modèle dans ce cas est $\boldsymbol{y} = \mathbf{X}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$. Montrer que la matrice $\mathbf{X}$ n'est plus de plein rang.

d) On peut pallier l'inconvénient remarqué en c) en imposant une contrainte sur les paramètres. Considérez, par exemple, la contrainte $\Sigma\alpha_i = 0$. Deux des 3 α_i suffisent alors pour exprimer toutes les moyennes, le 3^e pouvant toujours s'exprimer en fonction des deux autres. Montrer que le modèle peut s'écrire comme ceci : $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ où $\boldsymbol{\beta}' = [\beta_1 ; \beta_2 ; \beta_3] = [\mu ; \alpha_1 ; \alpha_2]$. [*Dans cette paramétrisation,* μ *est la moyenne globale* (*moyenne des* 3 *moyennes* μ_1, μ_2, μ_3), α_1 *et* α_2 *sont les écarts des moyennes* μ_1 *et* μ_2 *par rapport à cette moyenne globale*]

   (i) Identifiez la matrice $\mathbf{X}$.
   (ii) Exprimez H_o en termes de μ, α_1 et α_2.
   (iii) Exprimez H_o sous la forme $\mathbf{L}'\boldsymbol{\beta} = \mathbf{0}$.

e) Remplacez la contrainte $\Sigma\alpha_i = 0$ par la contrainte $\alpha_1 = 0$ et refaites l'exercice en c). Montrer que le modèle peut s'écrire comme ceci : $\boldsymbol{y} = \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ où $\boldsymbol{\gamma}' = [\gamma_1 ; \gamma_2 ; \gamma_3] = [\mu ; \alpha_2 ; \alpha_3]$. [*Dans cette paramétrisation,* μ *est la moyenne du premier groupe* α_2 *et* α_3 *sont les écarts des moyennes* μ_2 *et* μ_3 *par rapport à la moyenne du premier groupe*]

   (i) Identifiez la matrice $\mathbf{X}$.
   (ii) Exprimez H_o en termes de μ, α_2 et α_3
   (iii) Exprimez H_o sous la forme $\mathbf{L}'\boldsymbol{\gamma} = \mathbf{0}$.

f) Les vecteurs $\boldsymbol{\mu}$, $\boldsymbol{\beta}$ et $\boldsymbol{\gamma}$, trois paramétrisations parmi d'autres, ne sont que des transformations linéaires les unes des autres : c'est-à-dire, il existe des matrices $\mathbf{A}$ et $\mathbf{B}$ telles que $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\beta}$ et $\boldsymbol{\mu} = \mathbf{B}\boldsymbol{\gamma}$. Donc $\mathbf{X}\boldsymbol{\mu} = \mathbf{XAA}^{-1}\boldsymbol{\mu} = (\mathbf{XA})\boldsymbol{\beta}$ et $\mathbf{X}\boldsymbol{\mu} = \mathbf{XBB}^{-1}\boldsymbol{\mu} = (\mathbf{XB})\boldsymbol{\gamma}$. Identifiez $\mathbf{A}$ et $\mathbf{B}$,

3.2 Dans un modèle linéaire général $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ avec un vecteur $\boldsymbol{\beta} = [\beta_1 ; \beta_2 ; \beta_3 ; \beta_4]'$, considérer l'hypothèse

$$H_o : \beta_2 = \beta_3 = \beta_4$$

a) Cette hypothèse s'exprime aussi bien par $\mathbf{L}'\boldsymbol{\beta} = \mathbf{0}$, où $\mathbf{L}' = \begin{bmatrix} 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ que par $\mathbf{M}'\boldsymbol{\beta} = \mathbf{0}$, où $\mathbf{M}'$ = $\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$. En d'autres termes, une même hypothèse peut s'exprimer de différentes façons. Montrer, en général, que cela ne change pas la valeur de la statistique F pour tester une hypothèse: la valeur de F demeure la même, quel que soit le choix de la matrice.

b) Déterminer la matrice $\mathbf{T}$ telle que l'hypothèse H_o puisse être exprimée par $H_o : \boldsymbol{\beta} = \mathbf{T}\boldsymbol{\alpha}$ pour un certain vecteur $\boldsymbol{\alpha}$.

3.3 Soit un modèle linéaire général $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

a) Montrez que deux fonctions linéaires $\mathbf{L}_1'\hat{\boldsymbol\beta}$ et $\mathbf{L}_2'\hat{\boldsymbol\beta}$ sont indépendantes si et seulement si $\mathbf{L}_1'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}_2 = \mathbf{0}$.

b) Soit $\mathbf{L}_1'\hat{\boldsymbol\beta}$ et $\mathbf{L}_2'\hat{\boldsymbol\beta}$ deux fonctions linéaires indépendantes telles que $r([\mathbf{L}_1 \mid \mathbf{L}_2]) = r(\mathbf{X})$. Montrez que $(\mathbf{L}_1'\hat{\boldsymbol\beta})'[\mathbf{L}_1'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}_1]^{-1}\mathbf{L}_1'\hat{\boldsymbol\beta} + (\mathbf{L}_2'\hat{\boldsymbol\beta})'[\mathbf{L}_2'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}_2]^{-1}\mathbf{L}_2'\hat{\boldsymbol\beta} = \boldsymbol{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y}$ [*Vous pouvez admettre* (*ou démontrer, si vous préférez*) *le théorème suivant* : *Si* $\mathbf{A}_{n\times n}$ *et* $\mathbf{B}_{n\times n}$ *sont idempotentes et symétriques*, $\mathbf{AB} = \mathbf{0}$, *et* $r(\mathbf{A}+\mathbf{B}) = n$, alors $\mathbf{A}+\mathbf{B} = \mathbf{I}_n$].

3.4 Dans le problème présenté au numéro 1, considérez la paramétrisation $\boldsymbol{y} = \mathbf{X}\boldsymbol\mu + \boldsymbol\varepsilon$. H_o peut s'exprimer comme H_o : $\ell_1'\boldsymbol\mu = 0$ et $\ell_2'\boldsymbol\mu = 0$, où $\ell_1' = [1 ; -½ ; -½]$ **et** $\ell_2' = [0 ; 1 ; -1]$. Donc H_o : $\mathbf{L}'\boldsymbol\mu = \mathbf{0}$, où $\mathbf{L}' = \begin{bmatrix}\ell_1' \\ \ell_2'\end{bmatrix} = \begin{bmatrix}1 & -1/2 & -1/2 \\ 0 & 1 & -1\end{bmatrix}$.

a) Montrer que $\ell_1'\hat{\boldsymbol\mu}$ **et** $\ell_2'\hat{\boldsymbol\mu}$ ne sont pas indépendantes.

b) Montrez que $\ell_1'\hat{\boldsymbol\mu}$ **et** $\ell_2'\hat{\boldsymbol\mu}$ auraient été indépendantes si les effectifs des trois classes avaient été égaux.

c) De plus, montrez que si les effectifs des classes avaient été égaux, la somme des carrés expliquée SCE aurait été égale à SCE_1+SCE_2, SCE_i étant la somme des carrés du numérateur pour tester $\ell_i'\boldsymbol\mu = 0$.

d) Vérifiez numériquement que la propriété énoncée en c) n'est pas vérifiée dans le cas présenté.

3.5 Définissez le cosinus $\cos[\boldsymbol{x} ; \boldsymbol{y}]$ de l'angle entre deux vecteurs $\boldsymbol{x}$ et $\boldsymbol{y}$ par

$$\cos(\boldsymbol{x} ; \boldsymbol{y}) = \frac{\boldsymbol{x}'\boldsymbol{y}}{\sqrt{\boldsymbol{x}'\boldsymbol{x}}\sqrt{\boldsymbol{y}'\boldsymbol{y}}}$$

Montrez que $\cos^2[\boldsymbol{y} ; \mathbf{X}\boldsymbol\beta]$ est maximisé par rapport à $\boldsymbol\beta$ par $\hat{\boldsymbol\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y}$. Ce maximum n'est pas unique. Caractérisez l'ensemble des vecteurs $\tilde{\boldsymbol\beta}$ pour lesquels $\cos^2[\boldsymbol{y} ; \mathbf{X}\tilde{\boldsymbol\beta}]$ est également maximal. [*Utilisez l'inégalité de Cauchy-Schwarz. Notez que* $\mathbf{X}'\mathbf{X}$ *et son inverse sont définies positives et écrivez* $\boldsymbol{y}'\mathbf{X}\boldsymbol\beta = \boldsymbol{y}'\mathbf{HX}\boldsymbol\beta$]

3.6 Le coefficient de corrélation $r[\boldsymbol{u} ; \boldsymbol{v}]$ entre deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ est défini par $r[\boldsymbol{u} ; \boldsymbol{v}] = \cos[\mathbf{C}\boldsymbol{u} ; \mathbf{C}\boldsymbol{v}]$, où $\mathbf{C} = \mathbf{I}-\boldsymbol{ee}'/n$, $\boldsymbol{e}' = (1, \ldots , 1)$.

a) Supposez que $\boldsymbol{e} \in \mathcal{C}(\mathbf{X})$. Montrez que $r^2[\boldsymbol{y} ; \mathbf{X}\boldsymbol\beta]$ est aussi maximisé par $\hat{\boldsymbol\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y}$.

b) Le coefficient de corrélation multiple R entre $\boldsymbol{y}$ et les colonnes de $\mathbf{X}$ est défini par $R = r[\boldsymbol{y} ; \mathbf{X}\tilde{\boldsymbol\beta}]$, où $\tilde{\boldsymbol\beta}$ est une valeur de $\boldsymbol\beta$ qui maximise $r^2[\boldsymbol{y} ; \mathbf{X}\boldsymbol\beta]$. Montrez que

$$R^2 = \frac{\boldsymbol{y}'(\mathbf{P}-\mathbf{P}_o)\boldsymbol{y}}{\boldsymbol{y}'(\mathbf{I}-\mathbf{P}_o)\boldsymbol{y}} = \frac{\boldsymbol{y}'(\mathbf{I}-\mathbf{P}_o)\boldsymbol{y}-\boldsymbol{y}'(\mathbf{I}-\mathbf{P})\boldsymbol{y}}{\boldsymbol{y}'(\mathbf{I}-\mathbf{P}_o)\boldsymbol{y}}$$

où $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ et $\mathbf{P}_o = \boldsymbol{ee}'/n$. Ce résultat montre que R^2 représente une réduction relative dans une somme de carrés résiduelle: $\boldsymbol{y}'(\mathbf{I} - \mathbf{P}_o)\boldsymbol{y}$ est une somme de carrés résiduelle dans un modèle avec $\boldsymbol{e}$ comme matrice de design, alors que $\boldsymbol{y}'(\mathbf{I} - \mathbf{P})\boldsymbol{y}$ est la somme de carrés résiduelle dans le modèle avec toutes les variables indépendantes présentes.

3.7 Soit $\boldsymbol{y} = \mathbf{X}\boldsymbol\beta + \boldsymbol\varepsilon$ le modèle linéaire général, soit la partition suivante de $\boldsymbol{X}$:

$$\boldsymbol{X} = (\boldsymbol{x}_o ; \boldsymbol{x}_1 ; \mathbf{X}_2) \text{ où } \boldsymbol{x}_o' = \boldsymbol{e}' = (1 ; \ldots ; 1)$$

Désignons par $\mathbf{P}_i$ le projecteur orthogonal sur $\mathcal{C}(\boldsymbol{X}_i)$: $\mathbf{P}_i = \mathbf{X}_i(\mathbf{X}_i'\boldsymbol{X}_i)^{-1}\mathbf{X}_i'$; et par $\mathbf{P}_{ij}$ le projecteur orthogonal sur $\mathcal{C}(\mathbf{X}_{ij}) = \mathcal{C}(\mathbf{X}_i \mid \mathbf{X}_j)$: $\mathbf{P}_{ij} = \mathbf{X}_{ij}(\mathbf{X}_{ij}'\mathbf{X}_{ij})^{-1}\mathbf{X}_{ij}'$. Nous désignons toujours par $\mathbf{P}$ le projecteur orthogonal sur $\mathcal{C}(\mathbf{X})$: $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Le *coefficient de corrélation partielle* $r_{y1.2}$ *entre* $\boldsymbol{y}$ et $\boldsymbol{x}_1$ *étant donné* $\mathbf{X}_2$ est défini par

$$r_{y1.2} = r[\boldsymbol{y}; \boldsymbol{x}_1 \mid \mathbf{X}_2] = \cos[(\mathbf{I} - \mathbf{P}_{02})\boldsymbol{y} ; (\mathbf{I} - \mathbf{P}_{02})\boldsymbol{x}_1].$$

[Une meilleure notation aurait été $r_{y1.2} = r[\boldsymbol{y} ; \boldsymbol{x}_1 \mid (\boldsymbol{x}_o ; \mathbf{X}_2)]$, qui serait alors le cosinus de l'angle entre les projections des vecteurs $\boldsymbol{y}$ et $\boldsymbol{x}_1$ sur le complément orthogonal de $\mathcal{C}(\boldsymbol{x}_o \mid \mathbf{X}_2)$. Le coefficient de corrélation ordinaire serait alors le cas particulier $r = r[\boldsymbol{y} ; \boldsymbol{x}_1 \mid \boldsymbol{x}_o]$].

a) Montrez que $r_{y1.2} = r[(\mathbf{I} - \mathbf{P}_{02})\boldsymbol{y} ; (\mathbf{I} - \mathbf{P}_{02})\boldsymbol{x}_1]$.

b) Montrez que $r_{y1.2}^2 = [\boldsymbol{y}'(\mathbf{P} - \mathbf{P}_{02})\boldsymbol{y}]/[\boldsymbol{y}'(\mathbf{I} - \mathbf{P}_{02})\boldsymbol{y}]$.

Ce résultat montre que le coefficient de corrélation partielle représente la réduction relative de la somme de carrés résiduelle due à l'introduction de la variable 1 dans un modèle qui comprend déjà les variables contenues dans $\mathbf{X}_2$.

3.8 Dans le modèle $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ avec $\mathbf{X}_{n\times q} = (\boldsymbol{x}_o \,;\, \mathbf{X}_1)$ où $\boldsymbol{x}_o = \boldsymbol{e}$ et $\boldsymbol{\beta}' = (\beta_o \,;\, \boldsymbol{\beta}_1')$, montrez que la statistique F pour tester l'hypothèse $\boldsymbol{\beta}_1 = \mathbf{0}$ est une fonction croissante de R^2, le carré du coefficient de corrélation multiple [*Il suffit de montrer que* $F = \dfrac{n-q}{q-1} \times \dfrac{R^2}{1-R^2}$.]

3.9 Dans le modèle $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ avec $\mathbf{X}_{n\times q} = (\boldsymbol{x}_o | \boldsymbol{x}_1 | \mathbf{X}_2)$ où $\boldsymbol{x}_o = \boldsymbol{e}$ et $\boldsymbol{\beta}' = (\beta_o \,;\, \beta_1 \,;\, \boldsymbol{\beta}_2')$, montrez que la statistique F pour tester l'hypothèse que $\beta_1 = 0$ est une fonction croissante de $r^2_{y1.2}$, le carré du coefficient de corrélation partielle entre $\boldsymbol{y}$ et $\boldsymbol{x}_1$ étant donné $\mathbf{X}_2$, soit $F = \dfrac{n-q}{q-1} \cdot \dfrac{r^2_{i1.2}}{1-r^2_{i1.2}}$

3.10 Considérer le modèle linéaire général $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, avec les suppositions usuelles, mais dans lequel le paramètre $\boldsymbol{\beta}$ est contraint par $\boldsymbol{\beta} = \mathbf{T}\boldsymbol{\gamma}+\boldsymbol{b}$, où $\mathbf{T}$ est une matrice fixe connue de plein rang et $\boldsymbol{b}$ est un vecteur connu fixe. Estimer $\boldsymbol{\beta}$ sous cette contrainte. [$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}(\mathbf{T}\boldsymbol{\gamma} + \boldsymbol{b}) = \mathbf{X}\mathbf{T}\boldsymbol{\gamma} + \mathbf{X}\boldsymbol{b} \Leftrightarrow \boldsymbol{y}\text{-}\mathbf{X}\boldsymbol{b} = \mathbf{X}\mathbf{T}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$. *À partir d'ici on utilise les résultats connus dans un modèle linéaire dont le vecteur des observations est* $\boldsymbol{y}$-$\mathbf{X}\boldsymbol{b}$ *et la matrice de design est* $\mathbf{XT}$.]

3.11 Considérer le modèle linéaire général $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, avec les suppositions usuelles, mais dans lequel le paramètre $\boldsymbol{\beta}$ est contraint par $\mathbf{L}'\boldsymbol{\beta} = \boldsymbol{b}$, où $\mathbf{L}$ est une matrice fixe connue de rang plein et $\boldsymbol{b}$ est un vecteur fixe. Estimer $\boldsymbol{\beta}$ sous cette contrainte. [*On peut supposer sans perte de généralité que les colonnes de* $\mathbf{L}$ *sont orthonormales* (*pourquoi ?*) *On peut écrire* $\boldsymbol{\beta} = \mathbf{L}\boldsymbol{\alpha} + \mathbf{T}\boldsymbol{\gamma}$, *où* $\mathbf{T}$ *est une matrice dont les colonnes forment une base orthonormale de* $\mathcal{C}(\mathbf{L})$. *Montrer que* $\boldsymbol{\beta} = \mathbf{L}(\mathbf{L}'\mathbf{L})^{-1}\boldsymbol{b} + \mathbf{T}\boldsymbol{\gamma}$ *et considérez* $\boldsymbol{y} - \mathbf{X}\mathbf{L}(\mathbf{L}'\mathbf{L})^{-1}\boldsymbol{b}$ *comme votre vecteur d'observations*].

3.12 Au numéro précédent, si $\mathbf{A}$ est $q\times r$ de rang r, il existe une sous matrice carrée inversible. Supposons que $\mathbf{A} = [\mathbf{A}_1 \,|\, \mathbf{A}_2]$, où $\mathbf{A}_1$ est $r\times r$ et non singulière. Partitionnons $\mathbf{X} = [\mathbf{X}_1 \,|\, \mathbf{X}_2]$ et $\boldsymbol{\beta} = \begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix}$ de façon conforme ($\mathbf{X}_1$ ayant r colonnes et $\boldsymbol{\beta}_1$ r composantes). La contrainte peut s'écrire comme $\mathbf{A}_1\boldsymbol{\beta}_1+\mathbf{A}_2\boldsymbol{\beta}_2 = \boldsymbol{b}$. Incorporer cette contrainte dans l'écriture du modèle $\boldsymbol{y} = [\mathbf{X}_1 \,|\, \mathbf{X}_2]\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix} + \boldsymbol{\varepsilon}$ pour montrer que l'estimateur de $\boldsymbol{\beta}$ peut s'écrire comme $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\boldsymbol{\beta}}_1\\ \hat{\boldsymbol{\beta}}_2\end{bmatrix}$, où

$\hat{\boldsymbol{\beta}}_2 = (\mathbf{X}_r'\mathbf{X}_r)^{-1}\mathbf{X}_r'\boldsymbol{y}_r$, $\mathbf{X}_r = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2$, $\boldsymbol{y}_r = \boldsymbol{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\boldsymbol{b}$.

3.13 Un importateur reçoit des oranges de 5 pays différents, A, B, C, D, et E. Pour estimer les poids moyens μ_1, μ_2, μ_3, μ_4 et μ_5 de ces pays, on procède de la façon suivante : On place en même temps 3 oranges, provenant de trois pays différents, et on note le poids total. On effectue 10 pesées (toutes les combinaisons possibles de 3 oranges). Voici les résultats :

Poids total	Pays
$y_1 = 552$	A ; B ; C
$y_2 = 560$	A ; B ; D
$y_3 = 592$	A ; B ; E
$y_4 = 564$	A ; C ; D
$y_5 = 575$	A ; C ; E
$y_6 = 575$	A ; D ; E
$y_7 = 571$	B ; C ; D
$y_8 = 587$	B ; C ; E
$y_9 = 589$	B ; D ; E
$y_{10} = 601$	C ; D ; E

Les poids des oranges sont de loi normale, de variance σ^2 dans chaque population. L'observation y_1 dans le tableau ci-dessus, par exemple, est une somme de trois variables aléatoires, $y_1 = z_1 + z_2 + z_3$, de moyennes μ_1, μ_2 et μ_3, respectivement, et $\mathrm{Var}(z_1) = \mathrm{Var}(z_2) = \mathrm{Var}(z_3) = \sigma^2$ ($\Rightarrow \mathrm{Var}(y_1) = 3\sigma^2$.)

a) Estimer le vecteur des moyennes $\boldsymbol{\mu}' = [\mu_1 \,;\, \mu_2 \,;\, \mu_3 \,;\, \mu_4 \,;\, \mu_5]$.

b) Estimer la matrice de covariance de l'estimateur $\hat{\boldsymbol{\mu}}$ de $\boldsymbol{\mu}$.

c) Estimer le vecteur des contrastes [μ_2-μ_1 ; μ_3-μ_2 ; μ_4-μ_3 ; μ_5-μ_4] et déterminer la matrice de covariance de son estimateur.

d) Déterminer un intervalle de confiance à 95 % pour μ_3-μ_2.

e) Tester l'hypothèse que $\mu_1 = \mu_2$.

f) Tester l'hypothèse que les 5 moyennes sont égales.

g) On aura pesé en tout 30 oranges en 10 pesées. Est-ce qu'on aurait mieux fait de peser deux fois 3 oranges d'un même pays simultanément, c'est-à-dire,

- une pesée de 3 oranges de A dans la balance, puis une pesée de 3 oranges de B,…, finalement une pesée de 3 oranges de E (5 pesées, 15 oranges);
- et on recommence une deuxième fois (encore 5 pesées, encore 3 oranges).

Déterminer les matrices de covariance dans les deux modèles et montrer qu'une certaine différence est définie positive.

3.14 Considérer le modèle $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ avec $\mathbf{X}_{n\times q} = (\boldsymbol{x}_1 \mid \mathbf{X}_2)$ où $\boldsymbol{\beta} = \begin{bmatrix}\beta_1\\ \boldsymbol{\beta}_2\end{bmatrix}$. Soit $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_1\\ \hat{\boldsymbol{\beta}}_2\end{bmatrix}$ l'estimateur habituel de $\boldsymbol{\beta}$.

Montrez que $\mathrm{Var}(\hat{\beta}_1) \geq \dfrac{\sigma^2}{\boldsymbol{x}_1'\boldsymbol{x}_1}$; et que le minimum est atteint lorsque les colonnes de $\mathbf{X}_2$ sont orthogonales à $\boldsymbol{x}_1$.

[Déterminer la première composante de la matrice de covariance $\sigma^2\begin{bmatrix}\boldsymbol{x}_1'\boldsymbol{x}_1 & \boldsymbol{x}_1'\mathbf{X}_2\\ \mathbf{X}_2'\boldsymbol{x}_1 & \mathbf{X}_2'\mathbf{X}_2\end{bmatrix}^{-1}$ de $\hat{\boldsymbol{\beta}}$]

3.15 Considérer le modèle $\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ avec $\mathbf{X}_{n\times q} = (\boldsymbol{x}_o \mid \boldsymbol{x}_1 \mid \mathbf{X}_2)$ où $\boldsymbol{x}_o = \boldsymbol{e}$ et $\boldsymbol{\beta} = \begin{bmatrix}\beta_o\\ \beta_1\\ \boldsymbol{\beta}_2\end{bmatrix}$. Soit $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_o\\ \hat{\beta}_1\\ \hat{\boldsymbol{\beta}}_2\end{bmatrix}$ l'estimateur habituel de $\boldsymbol{\beta}$. Montrez que $\mathrm{Var}(\hat{\beta}_1) \geq \dfrac{\sigma^2}{\boldsymbol{x}_1'\mathbf{C}\boldsymbol{x}_1}$, où $\mathbf{C} = \mathbf{I} - \boldsymbol{e}(\boldsymbol{e}'\boldsymbol{e})^{-1}\boldsymbol{e}'$; et que le minimum est atteint lorsque toutes les covariances entre $\boldsymbol{x}_1$ et les autres variables dans $\mathbf{X}_2$ sont nulles.

3.16 On doit peser 4 objets le plus précisément possible dans une balance à deux plateaux. On pourrait peser chaque objet deux fois (et prendre ensuite la moyenne). Ou encore, on pourrait peser 8 fois les 4 objets à la fois, de la façon suivante: on place certains des objets sur un des plateaux, les autres dans l'autre, et on utilise des poids pour équilibrer. Soit μ_1, μ_2 μ_3 et μ_4 les poids *réels* des 4 objets. Chaque résultat y est une somme d'espérance $w_1\mu_1+w_2\mu_2+w_3\mu_3+w_4\mu_4$, où les w_i prennent les valeurs 1 ou -1 selon que l'objet de moyenne μ_i est dans le plateau de gauche ou de droite. Voici les résultats de 8 pesées :

Objet 1	Objet 2	Objet 3	Objet 4	*y*
1	1	1	1	20,2
1	-1	1	-1	8,1
1	1	-1	-1	9,7
1	-1	-1	1	1,9
1	1	1	1	19,9
1	-1	1	-1	8,3
1	1	-1	-1	10,2
1	-1	-1	1	1,8

[Par exemple, l'observation y de la 2^{e} ligne a pour moyenne μ_1-μ_2+μ_3-μ_4]. Le modèle est donc $\boldsymbol{y} = \mathbf{X}\boldsymbol{\mu}+\boldsymbol{\varepsilon}$, où

$$\mathbf{X} = \begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\\ 1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}.]$$

On supposera que $\mathrm{Var}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$.

a) Estimer le vecteur $\boldsymbol{\mu}$' = [μ_1 ; μ_2 ; μ_3 ; μ_4] par la méthode des moindres carrés et estimer la matrice de covariance de l'estimateur.

b) Supposons que la variance σ^2 définie ci-dessus est la dispersion du résultat d'une pesée, quel que soit le nombre d'objets dans les balances. Comparez la variance $\mathrm{Var}(\hat{\mu}_i)$ dans le plan d'expérience décrit en a) avec une approche qui consiste à peser chaque objet séparément, deux fois.

c) Généralisation du problème : on a m objets à peser, on pèse différentes combinaisons n fois, $n \geq m$. Chaque pesée peut être représentée par une ligne d'une matrice **X** comme celle montrée plus haut. On considère l'ensemble de toutes les matrices de cette forme, dont les composantes sont des 1, des -1 et également des 0 (correspondant à un objet qui n'est *pas* mis dans la balance.) Montrez que pour minimiser la variance $\mathrm{Var}(\hat{\mu}_i)$, la matrice **X** ne doit comprendre aucun 0 et doit satisfaire la condition $\mathbf{X}'\mathbf{X} = n\mathbf{I}_m$.

3.17 L'ellipsoïde $(\hat{\boldsymbol\beta}-\boldsymbol\beta)'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol\beta}-\boldsymbol\beta) \leq q\hat{\sigma}^2 F_{q;n-q;\alpha}$ est une région de confiance de niveau 1-α pour le paramètre $\boldsymbol\beta$. Dans ce problème, on vous demande ce qu'on peut en déduire à propos d'une fonction linéaire $\boldsymbol\ell'(\hat{\boldsymbol\beta}-\boldsymbol\beta)$, en particulier, quelle est la borne supérieure a de $[\boldsymbol\ell'(\hat{\boldsymbol\beta}-\boldsymbol\beta)]^2$ qui découle de l'inégalité $(\hat{\boldsymbol\beta}-\boldsymbol\beta)'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol\beta}-\boldsymbol\beta) \leq q\hat{\sigma}^2 F_{q;n-q;\alpha}$.

a) Montrer que si $\boldsymbol\ell$ est un vecteur fixe, alors $(\boldsymbol\ell'\boldsymbol b)^2 \leq k(\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol\ell)$ pour tout vecteur $\boldsymbol b$ tel que $\boldsymbol b'\mathbf{X}'\mathbf{X}\boldsymbol b = k$, et que la borne supérieure est atteinte lorsque $\boldsymbol b = \sqrt{\dfrac{k}{\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol\ell}}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol\ell$ [*Suggestion* : $(\boldsymbol\ell'\boldsymbol b)^2 = [\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1/2}(\mathbf{X}'\mathbf{X})^{1/2}\boldsymbol b]^2$, puis utiliser l'inégalité de Cauchy-Schwartz pour montrer que $[\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1/2}(\mathbf{X}'\mathbf{X})^{1/2}\boldsymbol b]^2 \leq (\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol\ell)(\boldsymbol b'\mathbf{X}'\mathbf{X}\boldsymbol b)$].

[*Rappel* : Soit $\boldsymbol x$ et $\boldsymbol y$ deux vecteurs de même dimension. Alors $(\boldsymbol x'\boldsymbol y)^2 \leq (\boldsymbol x'\boldsymbol x)(\boldsymbol y\boldsymbol y)$. La borne est atteinte si $\boldsymbol x = \lambda\boldsymbol y$ pour un certain λ]

b) En déduire que $(\hat{\boldsymbol\beta}-\boldsymbol\beta)'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol\beta}-\boldsymbol\beta) \leq q\hat{\sigma}^2 F_{q;n-q;\alpha} \Rightarrow [\boldsymbol\ell'(\hat{\boldsymbol\beta}-\boldsymbol\beta)]^2 \leq \hat{\sigma}^2 q F_{q;n-q;\alpha}\boldsymbol\ell'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol\ell$.