Contrôle des connaissances du cours « Bruits et Signaux » de l'École Doctorale d'Astrophysique.

Corrigé

Le mercredi 15 décembre 1999 de 14h à 16h amphithéâtre de l'IAP.

Le contrôle est noté sur 20, le barème est indiqué à la fin de chaque question. Le polycopié du cours et les notes de cours sont les seuls documents autorisés.

Un modèle à 2 paramètres.

L'interprétation d'une observation a été faite à partir de l'ajustement des données par la méthode des moindres carrés. Le modèle qui a permis cet ajustement dépend de deux paramètres : θ_1 et θ_2 (voir à titre d'exemple la photocopie, jointe en annexe, d'une page du dernier *Astrophysical Journal Letters*). On notera $\boldsymbol{y}$ les données et $\boldsymbol{\mu}$ le modèle et on supposera que $\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{\epsilon}$, où $\boldsymbol{\epsilon}$ est un bruit normal (gaussien) de moyenne nulle. Les symboles $\boldsymbol{y}$, $\boldsymbol{\mu}$ et $\boldsymbol{\epsilon}$ représentent des vecteurs colonnes :

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad , \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \quad , \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad .$$

1. Que vaut la matrice des variances-covariances $\boldsymbol{V}$ du bruit $\boldsymbol{\epsilon}$ si on suppose que ce bruit n'est pas corrélé et que les erreurs sur les mesures y_i (composant le vecteur $\boldsymbol{y}$) ont toutes le même écart type σ ?
   [1 point]
   Réponse :
   $$\boldsymbol{V} = \sigma^2 \boldsymbol{I} \,, \quad \text{où } \boldsymbol{I} \text{ est la matrice identité.}$$
2. Même question pour le vecteur $\boldsymbol{y}$ représentant les mesures ?
   [1 point]
   Réponse :
   Pour la moyenne :
   $$\mathrm{E}\{\boldsymbol{y}\} = \mathrm{E}\{\boldsymbol{\mu} + \boldsymbol{\epsilon}\} = \boldsymbol{\mu} + \mathrm{E}\{\boldsymbol{\epsilon}\} = \boldsymbol{\mu} \,,$$
   et pour la matrice des variances-covariances :
   $$\mathrm{E}\{(\boldsymbol{y} - \boldsymbol{\mu})^t (\boldsymbol{y} - \boldsymbol{\mu})\} = \mathrm{E}\{\boldsymbol{\epsilon}^t \boldsymbol{\epsilon}\} = \boldsymbol{V} = \sigma^2 \boldsymbol{I} \,.$$

3. Les mesures y_i sont elles des variables aléatoires indépendantes ? Justifiez brièvement votre réponse.
   [1 point]
   Réponse :
   Oui, car des variables aléatoires normales et non corrélées sont indépendantes, c'est une propriété caractéristique des variables aléatoires normales.
4. On suppose de plus que le modèle $\boldsymbol{\mu}$ dépend linéairement des paramètres $\boldsymbol{\theta}$ et ceci de façon non singulière. Quelle est, au sens des moindres carrés, l'expression de la meilleure estimation $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$?
   [1 point]
   Réponse :
   C'est une simple question de cours :
$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t\boldsymbol{y} \quad \text{où} \quad \boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\theta}\,.$$
5. Peut-on dire que $\widehat{\boldsymbol{\theta}}$ est une combinaison linéaire des observations ?
   [1 point]
   Réponse :
   Oui, car on obtient $\widehat{\boldsymbol{\theta}}$ par $\widehat{\boldsymbol{\theta}} = \boldsymbol{T}\boldsymbol{y}$ où $\boldsymbol{T} = \left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t$ est une matrice rectangulaire (k,n) et où k est le nombre de paramètres (ici égal à deux) et n est le nombre de données.
6. Les valeurs estimées $\widehat{\theta}_1$ et $\widehat{\theta}_2$ des paramètres θ_1 et θ_2 sont des variables aléatoires car elles dépendent des $\boldsymbol{y}$ qui sont elles-mêmes des variables aléatoires. Quelles sont la moyenne et la matrice des variances-covariances $\boldsymbol{\Sigma}$ de $\widehat{\boldsymbol{\theta}}$?
   [2 points]
   Réponse :
$$\mathrm{E}\{\widehat{\boldsymbol{\theta}}\} = \boldsymbol{\theta} \quad \text{et} \quad \boldsymbol{\Sigma} = \sigma^2\left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}.$$
   On a :
$$\begin{aligned}
\boldsymbol{X}^t\boldsymbol{X} &= \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} \\ \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i2}^2 \end{pmatrix},
\end{aligned}$$
$$\Delta = \sum_{i=1}^n x_{i1}^2 \sum_{j=1}^n x_{j2}^2 - \big(\sum_{i=1}^n x_{i1}x_{i2}\big)^2,$$
$$\boldsymbol{\Sigma} = \sigma^2\left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1} = \frac{\sigma^2}{\Delta}\begin{pmatrix} \sum_{i=1}^n x_{i2}^2 & -\sum_{i=1}^n x_{i1}x_{i2} \\ -\sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i1}^2 \end{pmatrix}.$$
7. Quelle est la loi suivie par le couple $(\widehat{\theta}_1, \widehat{\theta}_2)$?
   [1 point]
   Réponse : Le vecteur $\widehat{\boldsymbol{\theta}} = \boldsymbol{T}\boldsymbol{y}$ est une combinaison linéaire de variables aléatoires normales, il est donc lui-même normal, on a :
$$\widehat{\boldsymbol{\theta}} = \mathcal{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma})\,.$$

8. Quelles sont les lois (marginales) suivies par $\widehat{\theta}_1$ et $\widehat{\theta}_2$?
   [1 point]
   Réponse :
   Ce sont aussi des lois normales (à une dimension cette fois-ci) :
$$\widehat{\theta}_1 = \mathcal{N}(\theta_1, \Sigma_{11})\,, \quad \widehat{\theta}_2 = \mathcal{N}(\theta_2, \Sigma_{22})\,.$$
9. Les variables $\widehat{\theta}_1$ et $\widehat{\theta}_2$ sont elles indépendantes ?
   [1 point]
   Réponse :
   En général non, il n'y a aucune raison pour que $\Sigma_{12} = 0$.
10. Si les erreurs de mesure $\boldsymbol{\epsilon}$ ne sont pas normales, à quelle(s) condition(s) peut-on dire que les $\widehat{\theta}_i$ suivent approximativement une loi normale.
    [2 points]
    Réponse :
    On a $\theta_i = \sum_{j=1}^{n} T_{ij} y_j$. Les θ_i sont donc la somme des variables aléatoires $y'_j = T_{ij} y_j = T_{ij}(\mu_j + \epsilon_j)$. Si les variables aléatoires ϵ_j possèdent une moyenne, sont deux à deux indépendantes et si, de plus, elles possèdent un moment absolu supérieur à 2 (condition de Liapounov) alors les θ_i convergent en loi vers la loi normale (il s'agit d'une version du théorème central limite).
11. On suppose de nouveau que les variables aléatoires ϵ_i suivent une loi normale, sont indépendantes et possèdent toutes le même écart type σ. Comme on le sait, les $\widehat{\theta}_i$ sont obtenus en cherchant le minimum de l'expression :
$$S(\boldsymbol{\theta}) = (\boldsymbol{y} - \boldsymbol{\mu})^t(\boldsymbol{y} - \boldsymbol{\mu}) = \sum_{i=1}^{n} (y_i - \mu_i)^2\,.$$
    Montrer par un calcul direct que le résidu d'ajustement : $\widehat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \widehat{\boldsymbol{\mu}}$ est orthogonal au sous espace des modèles. C'est-à-dire : $\boldsymbol{\mu}^t \widehat{\boldsymbol{\epsilon}} = 0$ quelque soit μ.
    [2 points]
    Réponse :
$$\begin{aligned}
\boldsymbol{\mu}^t \widehat{\boldsymbol{\epsilon}} &= (\boldsymbol{X}\boldsymbol{\theta})^t(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\theta}})\,, \\
&= \boldsymbol{\theta}^t \boldsymbol{X}^t (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t)\boldsymbol{y}\,, \\
&= \boldsymbol{\theta}^t \boldsymbol{X}^t \boldsymbol{y} - \boldsymbol{\theta}^t \boldsymbol{X}^t \boldsymbol{X}(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}\,, \\
&= 0\,.
\end{aligned}$$
12. En déduire que l'on peut écrire :
$$S(\widehat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta}) = S(\widehat{\boldsymbol{\theta}}) + \Delta\boldsymbol{\theta}^t \boldsymbol{X}^t \boldsymbol{X} \Delta\boldsymbol{\theta}\,.$$
    C'est-à-dire :
$$\frac{S(\widehat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta})}{\sigma^2} = \frac{S(\widehat{\boldsymbol{\theta}})}{\sigma^2} + \Delta\boldsymbol{\theta}^t \boldsymbol{\Sigma}^{-1} \Delta\boldsymbol{\theta}\,.$$
    On a posé : $\Delta\boldsymbol{\theta} = \boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}$.

[1 point]

Réponse :

Le calcul est immédiat :

$$
\begin{aligned}
S(\widehat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta}) &= [\boldsymbol{y} - \boldsymbol{X}(\widehat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta})]^t [\boldsymbol{y} - \boldsymbol{X}(\widehat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta})]\,, \\
&= [(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\theta}}) - \boldsymbol{X}\Delta\boldsymbol{\theta}]^t [(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\theta}}) - \boldsymbol{X}\Delta\boldsymbol{\theta}]\,, \\
&= [\widehat{\boldsymbol{\epsilon}} - \boldsymbol{X}\Delta\boldsymbol{\theta}]^t [\widehat{\boldsymbol{\epsilon}} - \boldsymbol{X}\Delta\boldsymbol{\theta}]\,, \\
&= \widehat{\boldsymbol{\epsilon}}^t\widehat{\boldsymbol{\epsilon}} - \underbrace{\epsilon^t \boldsymbol{X}\Delta\boldsymbol{\theta}}_{=0} - \underbrace{\Delta\boldsymbol{\theta}^t\boldsymbol{X}^t\widehat{\boldsymbol{\epsilon}}}_{=0} + \Delta\boldsymbol{\theta}^t\boldsymbol{X}^t\boldsymbol{X}\Delta\boldsymbol{\theta}\,,\,, \\
&= S(\widehat{\boldsymbol{\theta}}) + \Delta\boldsymbol{\theta}^t\boldsymbol{X}^t\boldsymbol{X}\Delta\boldsymbol{\theta}\,.
\end{aligned}
$$

On pose souvent :

$$\chi^2(\boldsymbol{\theta}) = \frac{S(\boldsymbol{\theta})}{\sigma^2}\,,$$

ce qui autorise l'écriture :

$$\Delta\chi^2 = \Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}\,.$$

13. Quelle est la loi suivie par la variable aléatoire $\Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$?

    [1 point]

    Réponse :

    Une loi du χ^2 à deux degrés de liberté (à k degrés de liberté en général). Voir cours sur les lois normales à plusieurs dimensions.

14. On cherche à délimiter une zone de confiance autour de $\widehat{\boldsymbol{\theta}}$ qui contienne le vrai paramètre $\boldsymbol{\theta}$ avec la probabilité γ. Dans un premier temps, on se place en $\boldsymbol{\theta}$ et on cherche à déterminer la zone où se situe $\widehat{\boldsymbol{\theta}}$ avec la probabilité γ.

    Si on délimite cette zone par une frontière où la densité de probabilité de $\widehat{\boldsymbol{\theta}}$ est constante, quelle est l'allure de cette zone ?

    [1 point]

    Réponse :

    La variable aléatoire $\widehat{\boldsymbol{\theta}}$ suivant une loi normale de moyenne $\boldsymbol{\theta}$ et de matrice des variances-covariances $\boldsymbol{\Sigma}$ sa densité de probabilité est proportionnelle à $\exp\{-\frac{1}{2}\chi^2\}$, avec $\chi^2 = \Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$. La zone où se trouve $\widehat{\boldsymbol{\theta}}$ est une ellipse centré sur $\boldsymbol{\theta}$ car l'équation $\chi^2 =$ constante est celle d'une ellipse.

15. Que vaut l'expression $\Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$ lorsque γ vaut respectivement : $0,68$; $0,90$ et $0,99$?

    [1 point]

    Réponse :

    Comme dit plus haut, la forme quadratique $\Delta\chi^2 = \Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$ suit une loi du χ^2 à deux degrés de liberté, par définition de la fonction de répartition, on a :

    $$\Pr\{\Delta\chi^2 \leq k^2\} = F_{\chi^2}(k^2)\,.$$

    Sur la frontière qui englobe la probabilité γ on a :

    $$\Pr\{\Delta\chi^2 \leq k_\gamma^2\} = F_{\chi^2}(k_\gamma^2) = \gamma\,,$$

soit $k_\gamma^2 = F_{\chi^2}^{-1}(\gamma)$. On obtient ici pour deux degrés de liberté :

$$\begin{aligned} k_{0,68}^2 &= (1.51317)^2 \approx 2,28 \\ k_{0,9}^2 &= (2.14597)^2 \approx 4,61 \\ k_{0,99}^2 &= (3.03485)^2 \approx 9,21 \end{aligned}$$

d'après la table 6.4 p. 103 du cours.

16. Quel est l'ensemble des $\boldsymbol{\theta}$ (zone de confiance) qui sont susceptibles d'englober $\widehat{\boldsymbol{\theta}}$ dans la zone d'erreur qui leur est associée ?

    [1 point]

    Réponse :

    Tous les $\boldsymbol{\theta}$ possibles sont ceux qui sont à la distance $\Delta\chi^2 \leq k_\gamma^2$. L'expression $\Delta\chi^2 = \Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$ est symétrique par rapport à $\boldsymbol{\theta}$ et $\widehat{\boldsymbol{\theta}}$, de plus $\boldsymbol{\Sigma}^{-1}$ est constant, par conséquent la zone de confiance est délimitée par la même ellipse mais centrée autour de $\widehat{\boldsymbol{\theta}}$.

17. Que vaut l'accroissement de $\frac{S(\boldsymbol{\theta})}{\sigma^2}$ sur la frontière de la zone de confiance (au niveau de confiance γ) par rapport à sa valeur minimum ?

    [1 point]

    Réponse :

    Nous avons vu que cet accroissement $\Delta\chi^2 = \frac{S(\boldsymbol{\theta})}{\sigma^2} - \frac{S(\widehat{\boldsymbol{\theta}})}{\sigma^2}$ est égal à $\Delta\boldsymbol{\theta}^t\boldsymbol{\Sigma}^{-1}\Delta\boldsymbol{\theta}$, il est donc identique aux valeurs calculées plus haut. Soit : $2,28$; $4,61$ et $9,21$ pour une confiance de $\gamma = 0,68$; $0,90$ et $0,99$ respectivement. Ce sont les valeurs que l'on trouve dans la légende de la Fig, 1 de l'article donné en exemple.