

**Examen DEA SICOM, Cours Identification**  
**Luc Pronzato, Décembre 2001**

**Problème 1 :**

On considère un problème d'estimation de paramètres pour un modèle décrit par

$$y_m(\theta, t) = \frac{\theta_1 + t}{\theta_2 t}$$

avec  $\theta = (\theta_1, \theta_2)^\top$  le vecteur des paramètres à estimer. On suppose que les erreurs de mesures  $\epsilon_i$ , variables aléatoires indépendantes et identiquement distribuées, suivent une loi normale de moyenne 0 et de variance  $\sigma^2$ . On observe

$$y_i = y_m(\bar{\theta}, t_i) + \epsilon_i, \quad i = 1, \dots, N,$$

avec  $t_i$  le  $i$ ème instant d'observation et  $\bar{\theta}$  la vraie valeur (inconnue) de  $\theta$ . On estime  $\theta$  par moindres carrés et on note  $\hat{\theta}$  la valeur estimée de  $\theta$  à partir des observations  $\mathbf{y} = (y_1, \dots, y_N)^\top$ .

**1.1)** Donner l'expression de la matrice d'information de Fisher  $\mathbf{M}$  et de la matrice de covariance (asymptotique) de  $\hat{\theta}$ .

**1.2)** On souhaite effectuer deux observations, aux instants  $t_1$  et  $t_2$ , avec

$$1 \leq t_1 \leq t_2 \leq 10,$$

et obtenir la plus grande précision possible sur l'estimation de  $\theta$ . On choisit pour cela les temps d'observation  $t_1$  et  $t_2$  qui maximisent le déterminant de  $\mathbf{M}$ .

**1.2.1)** Justifier ce choix.

**1.2.2)** Calculer  $\det \mathbf{M}$  et déterminer les valeurs optimales de  $t_1$  et  $t_2$  (penser à utiliser le changement de variable  $z = 1/t$ ).

**Problème 2 :**

On estime par moindres carrés les paramètres  $\theta$  d'un modèle

$$y_m(\theta, t) = \mathbf{r}^\top(t)\theta, \quad \theta \in \mathbb{R}^p.$$

Les erreurs de mesure  $\epsilon_i$  sont de moyenne nulle, non corrélées ( $\mathbf{E}\{\epsilon_i \epsilon_j\} = 0$  pour  $i \neq j$ ) de variance  $\mathbf{E}\{\epsilon_i^2\} = \sigma^2$  pour tout  $i$ .

**2.1)** Donner l'expression de l'estimateur  $\hat{\theta}$  en fonction du vecteur  $\mathbf{y}$  des observations,  $\mathbf{y} = (y_1, \dots, y_N)^\top$ ,  $N \geq p$ , et de la matrice  $\mathbf{R}$  des régresseurs (matrice dont la ligne  $i$  est donnée par  $\mathbf{r}^\top(t_i)$ ,  $i = 1, \dots, N$ ).

**2.2)** Calculer la moyenne et la matrice de covariance  $\mathbf{C}$  de  $\hat{\theta}$ .

**2.3)** On souhaite à présent prédire la valeur de la réponse du modèle à un instant  $t_0$  particulier. On utilise le prédicteur  $y_m(\hat{\theta}, t_0) = \mathbf{r}^\top(t_0)\hat{\theta}$ . Calculer la moyenne et la variance de la prédiction. Comment choisir les temps d'observation pour avoir la prédiction la plus précise possible?

**2.4)** On s'intéresse à présent aux prédictions obtenues pour tout  $t$  tel que  $\mathbf{r} = \mathbf{r}(t)$  est dans  $\mathcal{B}(\mathbf{0}, R)$ , la boule de centre 0 de rayon  $R$ . Si l'on souhaite obtenir la plus grande précision possible dans le pire cas, c'est-à-dire garantir une précision pour tout  $\mathbf{r}$  dans  $\mathcal{B}(\mathbf{0}, R)$ , comment choisir les temps d'observation? On exprimera le critère à optimiser en fonction des valeurs propres de  $\mathbf{C}$ .

Si l'on souhaite à présent obtenir la plus grande précision possible en moyenne, pour  $\mathbf{r}$  distribué suivant une loi normale de moyenne  $\mathbf{0}$  et de matrice de covariance  $\Omega$ , comment choisir les temps d'observation? On exprimera le critère à optimiser en fonction de  $\mathbf{C}$  et  $\Omega$ .

**Problème 3 :**

On souhaite estimer les paramètres d'un modèle dont la réponse à l'instant  $t$  est de la forme  $y_m(\theta, t)$ . On recueille  $N$  observations  $\mathbf{y} = (y_1, \dots, y_N)^\top$ , mais seulement  $p = \dim(\theta)$  instants  $t_i$  différents sont utilisés,  $i = 1, \dots, p$ . On note  $r_i$  le nombre d'observations recueillies à l'instant  $t_i$ , et  $y_{i_j}$  ces observations,  $j = 1, \dots, r_i$ , avec  $\sum_{i=1}^p r_i = N$ . On suppose que le vecteur  $\mathbf{y}_m(\theta) = [y_m(\theta, t_1), \dots, y_m(\theta, t_p)]^\top$  décrit  $\mathbb{R}^p$  quand  $\theta$  varie dans  $\mathbb{R}^p$ .

**3.1)** On utilise l'estimateur de norme  $L_1$ , non pondéré,

$$\hat{\theta}^1 = \arg \min_{\theta} \|\mathbf{y} - \mathbf{y}_m(\theta)\|_1 = \arg \min_{\theta} \sum_{i=1}^N |y_i - y_m(\theta, t_i)| = \arg \min_{\theta} \sum_{i=1}^p \sum_{j=1}^{r_i} |y_{i_j} - y_m(\theta, t_i)|.$$

On notera pour tout  $z \in \mathbb{R}^p$ ,  $z = (z_1, \dots, z_p)^\top$ ,

$$\mathcal{Y}_i^+ = \{y_{i_j}, j = 1, \dots, r_i, / z_i > y_{i_j}\}, \mathcal{Y}_i^- = \{y_{i_j}, j = 1, \dots, r_i, / z_i < y_{i_j}\},$$

et  $n_i^+, n_i^-$  le nombre d'éléments de  $\mathcal{Y}_i^+$  et  $\mathcal{Y}_i^-$ . Montrer que  $\hat{\theta}^1$  est solution d'un ensemble de  $p$  équations à  $p$  inconnues, l'équation numéro  $i$  faisant intervenir la médiane des  $y_{i_j}$ ,  $j = 1, \dots, r_i$ . S'aider d'un dessin peut être utile. On rappelle que pour  $r_i = 2k + 1$ ,  $\text{med}\{y_{i_j}\}$  est la  $k$ ème valeur des  $y_{i_j}$  classés par ordre croissant (ou décroissant), et, pour  $r_i = 2k$ ,  $\text{med}\{y_{i_j}\}$  prend une valeur quelconque entre les  $k$ ème et  $(k + 1)$ ème valeurs des  $y_{i_j}$  classés par ordre croissant (ou décroissant). Donner une condition suffisante portant sur les  $r_i$  pour avoir unicité de  $\hat{\theta}^1$ .

**3.2)** On utilise l'estimateur de norme  $L_\infty$ , non pondéré,

$$\hat{\theta}^\infty = \arg \min_{\theta} \|\mathbf{y} - \mathbf{y}_m(\theta)\|_\infty = \arg \min_{\theta} \max_{i=1, \dots, N} |y_i - y_m(\theta, t_i)| = \arg \min_{\theta} \max_{i=1, \dots, p} \max_{j=1, \dots, r_i} |y_{i_j} - y_m(\theta, t_i)|.$$

Montrer que  $\hat{\theta}^\infty$  est solution d'un ensemble de  $p$  équations à  $p$  inconnues, l'équation numéro  $i$  faisant intervenir  $(\max_j y_{i_j} + \min_j y_{i_j})/2$ .

**3.3)** Calculer  $\hat{\theta}^1$  et  $\hat{\theta}^\infty$  lorsque  $y_m(\theta, t) = \theta_1 + t\theta_2$ , avec les observations 1.0, 2, 1.25 obtenues en  $t = 0$  et 12.5, 11.25, 11.5 obtenues en  $t = 1$ .