

### Problème DEA 94-95, identification. Cours de Luc Pronzato

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

$$y_m(\mathbf{p}, x) = \frac{p_1 + x}{xp_2},$$

où  $\mathbf{p} = (p_1, p_2)^T$  désigne le vecteur de paramètres à estimer. On doit effectuer pour cela deux mesures, sous les conditions expérimentales respectives  $x_a$  et  $x_b$ . Les erreurs de mesures sont supposées indépendantes et de distribution normale de moyenne nulle et de variance  $\sigma^2$ . Il s'agit de choisir  $x_a$  et  $x_b$  de manière à assurer une estimation de  $\mathbf{p}$  la plus précise possible au sens du critère de  $D$ -optimalité (maximisation du déterminant de la matrice d'information de Fisher  $F$ ). On suppose que toutes les variables  $p_1, p_2, x_a, x_b$  sont strictement positives.

Calculer la matrice  $F$  en fonction de  $p_1, p_2, x_a, x_b$  et  $\sigma^2$ . Montrer que la valeur de  $\sigma^2$  est sans conséquence sur le choix optimal de  $x_a$  et  $x_b$ . On posera donc dans la suite  $\sigma^2 = 1$ .

Calculer le déterminant de  $F$  (il s'écrit simplement en fonction de  $x_a^2, x_b^2, (x_a - x_b)^2, p_2^6$ ).

On suppose que la variable expérimentale  $x$  satisfait les contraintes  $1 \leq x \leq 2$ . Calculer les deux valeurs de  $x$ , soit  $x_a$  et  $x_b$ , avec  $x_a < x_b$ , qui maximisent le déterminant de  $F$  (on pourra utiliser le changement de variable  $z = 1/x$ ).

On considère à présent une approche du type planification approximative, pour laquelle on suppose que 50% des observations seront effectuées pour  $x = x_a$  et 50% pour  $x = x_b$ ,  $x_a$  et  $x_b$  prenant les valeurs trouvées précédemment. On souhaite établir la  $D$ -optimalité de ce protocole expérimental  $\xi$ . Écrire la matrice d'information de Fisher moyenne par échantillon  $F_m(\xi)$ , puis calculer la fonction  $f(x) = d(x, \xi) = \mathbf{s}(x)^T F_m^{-1}(\xi) \mathbf{s}(x)$  avec  $\mathbf{s}(x)$  le vecteur de sensibilité de la réponse du modèle par rapport aux paramètres. Quelle est la valeur maximale de  $f(x)$  sur l'intervalle  $[1, 2]$ ? En quels points est-elle atteinte? Conclure quant à la  $D$ -optimalité du protocole  $\xi$ .