STATISTIQUES 2 / HPC - BIG DATA 2023

Modèle linéaire gaussien : estimateurs par maximisation de la vraisemblance

1)
La densité normale multidimensionnelle a pour expression :

$$f(Y)=\frac{1}{\sqrt{(2\pi)^n \det(\Gamma)}}\exp\left(-\frac{1}{2}{}^t(Y-\mu)\Gamma^{-1}(Y-\mu)\right)$$

Dans le cadre théorique du modèle linéaire gaussien on a : $Y=T\beta+e$ avec $e \sim N_n(0,\sigma^2 I_n)$,

$\Gamma=V[Y]=\sigma^2 I_n$ et $\mu=E[Y]=T\beta$ (T matrice $n \times q$ de rang q)

On obtient alors la vraisemblance normale sphérique de l'échantillon constitué par les n réalisations des composantes de Y, une fonction de q+1 variables (σ^2 et les q composantes du vecteur β) :

$$L(\beta,\sigma^2)=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\|Y-T\beta\|^2\right)$$

2)
Maximisons la Log-vraisemblance ln(L) , il vient :

$$\frac{\partial \ln L}{\partial \beta}=\frac{-1}{2\sigma^2}\frac{\partial}{\partial \beta}\|Y-T\beta\|^2 \quad \text{et} \quad \frac{\partial \ln L}{\partial \sigma^2}=\frac{-n}{2\sigma^2}+\frac{1}{2\sigma^4}\|Y-T\beta\|^2$$

On remarque que maximiser ln(L) revient à minimiser la norme au carré du vecteur $Y-T\beta$, les deux méthodes d'estimation (maximisation de vraisemblance et minimisation des moindres carrés) mènent donc ici rigoureusement au même estimateur du vecteur β.

On obtient donc : $\hat{\beta}_{MV}=\hat{\beta}_{MCO}$ puis $\hat{\sigma}^2_{MV}=\frac{1}{n}\left\|Y-T\hat{\beta}_{MV}\right\|^2$

Sinon, il fallait exprimer la dérivée selon β comme dans le cours. Je détaille ici les développements pour ceux qui n'auraient pas l'habitude de la dérivation matricielle :

$$\|Y-T\beta\|^2={}^tYY-{}^tYT\beta-{}^t(T\beta)Y+{}^t(T\beta)T\beta={}^tYY-2{}^t\beta{}^tTY+{}^t\beta{}^tTT\beta$$

$$\frac{\partial}{\partial \beta}\|Y-T\beta\|^2=2{}^tTT\beta-2{}^tTY$$

On obtient donc : $\hat{\beta}_{MV}=\hat{\beta}_{MCO}=({}^tTT)^{-1}\,{}^tTY$

----

3)
L'estimateur non biaisé de σ^2 obtenu en cours est : $\hat{\sigma}^2 = \frac{1}{n-q}\left\|Y - T\hat{\beta}_{MCO}\right\|^2$

$$\hat{\sigma}^2_{MV} = \frac{1}{n}\left\|Y - T\hat{\beta}_{MV}\right\|^2 = \frac{1}{n}\left\|Y - T\hat{\beta}_{MCO}\right\|^2 = \frac{n-q}{n}\hat{\sigma}^2$$

$$E[\hat{\sigma}^2_{MV}] = \frac{n-q}{n}E[\hat{\sigma}^2] = \frac{n-q}{n}\sigma^2 \quad \rightarrow \quad BIAIS[\hat{\sigma}^2_{MV}] = E[\hat{\sigma}^2_{MV}] - \sigma^2 = \frac{-q}{n}\sigma^2$$

L'estimateur de σ^2 obtenu par maximisation de la vraisemblance est donc biaisé, mais on note qu'il est asymptotiquement non biaisé.