28 nov. – 1 déc. 2005 Formation INRA ACTA ICTA

La Rochelle

Estimation des paramètres des modèles
« Principes généraux »

David Makowski

UMR Agronomie INRA/INA-PG

makowski@grignon.inra.fr

Mon cours sur l'estimation des paramètres comportent deux parties. La première partie présentent des principes généraux. La deuxième partie porte sur un problème particulier que je définirai par la suite.

Paramètres

$$f\,(x;\theta)$$

Les équations | Les variables d'entrée | **<u>Les paramètres</u>**

« ***Un paramètre*** *est une valeur numérique qui n'est pas calculé par le modèle et qui n'est pas une variable d'entrée mesurée ou observée* »

Tout d'abord, je définis ici ce qu'est un paramètre.

Un modèle comporte plusieurs éléments: une fonction, des variables d'entrée et des paramètres.

Estimation des paramètres

« consiste à approcher les valeurs des paramètres à partir *de données expérimentales* et/ou *d'informations issues de l'expertise* »

C'est important car

« Les *performances d'un modèle* vont dépendre de la méthode utilisée pour estimer les paramètres »

Trois problèmes d'estimation de complexité croissante

Pb.1: Modèle linéaire avec un seul paramètre.

Pb.2 : Modèle linéaire avec 2 paramètres.

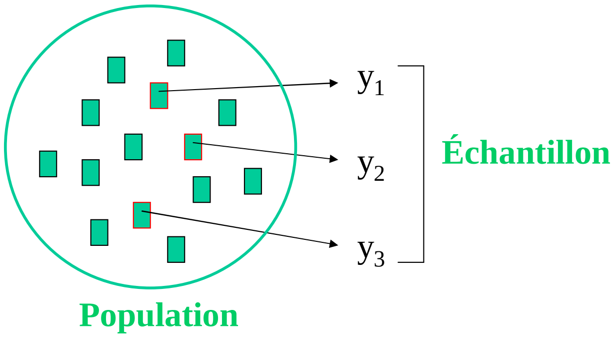
Pb.3 : Modèle non linéaire avec 18 paramètres.

Je vais aborder le problème de l'estimation des paramètres à travers trois problèmes de complexité croissante.

Estimation des paramètres des modèles

Problème 1

« Estimer le rendement moyen du colza en 2004 dans une petite région à partir de 3 mesures de rendement obtenues sur 3 parcelles »



Commençons par le problème le plus simple.

L'objectif est …

On considère donc un échantillon de trois mesures de rendement y1, y2, y3. Cet échantillon a été obtenu en réalisant des mesures sur trois parcelles agricoles tirées de façon aléatoire dans la région d'intérêt.

L'ensemble des parcelles de colza de la région représente ce qu'on appelle une population.

Quels paramètres doit-on estimer ?

Un seul paramètre à estimer, le rendement moyen de la région noté θ.

Estimer un ou plusieurs paramètres revient toujours à se poser une série de questions.

La première question que l'on se pose est …

Estimation des paramètres des modèles

Quelle information utiliser ?

Information disponible: un ***échantillon*** de trois mesures obtenues sur 3 parcelles de la ***population*** d'intérêt.

La deuxième

Estimation des paramètres des modèles

Quelle méthode d'estimation ?

Un estimateur du rendement de la parcelle est :

$$\hat{\theta} = \frac{y_1 + y_2 + y_3}{3}$$

Exemple :

• Si y_1=30, y_2=39 et y_3=35, la valeur estimée du rendement moyen est 34.7 q/ha.

• Si y_1=32, y_2=38 et y_3=39, la valeur estimée du rendement moyen est 36.3 q/ha.

« Un estimateur est une fonction qui relie le paramètre à des observations »

La troisième question est Quelle méthode d'estimation ?

Une solution naturelle consiste ici à faire simplement la moyenne des trois observations.

J'en profite ici pour définir ce qu'est un estimateur. C'est une fonction qui relie le paramètre aux observations. En changeant les observations, on change la valeur du paramètre en utilisant l'estimateur.

Voici un exemple. Supposons qu'on ait observé les valeurs suivantes…

Ici l'estimateur a une expression très simple. En pratique, ce sera souvent plus compliqué comme on le verra par la suite.

Estimation des paramètres des modèles

Cet estimateur est-il précis ?

$$E\left[\left(\hat{\theta}-\theta\right)^2\right]=\left[E\left(\hat{\theta}\right)-\theta\right]^2+\mathrm{var}\left(\hat{\theta}\right)$$

Erreur quadratique moyenne **Biais²** **Variance**

La dernière question consiste à se demander si l'estimateur est précis.

Pour étudier la précision d'un estimateur, il est utile de considérer l'erreur quadratique moyenne.

Theta est la vraie valeur du paramètre (inconnue), THETA^est la valeur estimée pour un échantillon de donnée, l'espérance est prise sur l'ensemble des échantillons de données possibles.

L'EQM est égale à la somme de deux termes. Le premier est le biais, c'est l'erreur systématique. Le biais indique si l'estimateur surestime ou sousestime systématiquement la vraie valeur du paramètre.

Le deuxième terme est la variance de l'estimateur. La variance donne une information sur la variabilité de THETA^ lorsqu'on change d'échantillon.

En pratique, on en connaît pas la vraie valeur theta, ni la vraie valeur du biais et de la variance de l'estimateur. Par contre, on peut approcher ces valeurs de différentes façons.

Estimation des paramètres des modèles

Cet estimateur est-il précis ?

a. Aspect théorique

« Sous certaines conditions, notre estimateur est ***sans biais*** et de ***variance minimale*** parmi les estimateurs sans biais »

Certains aspects théoriques peuvent être considérés pour juger de la précision d'un estimateur. Ainsi, …

Cet estimateur est-il précis ?

b. Variance de l'estimateur

On peut estimer $\mathrm{var}\left(\hat{\theta}\right)$ à partir des données

Exemple :

• Si y_1=30, y_2=39 et y_3=35, la valeur estimée de la variance est 6.78 q²/ha², soit e.t=2.6 q/ha.

• Si y_1=32, y_2=38 et y_3=39, la valeur estimée de la variance est 4.78 q²/ha², soit e.t=2.19 q/ha.

On peut également estimer la variance de l'estimateur à partir des données.

Problème 2

« Estimer les paramètres du modèle $f(x;\theta_1,\theta_2)$ »

$$f(x;\theta_1,\theta_2) = \theta_1 + \theta_2\, x$$

Azote absorbé par le colza

Dose d'engrais

Le modèle simule l'azote absorbé en fonction de la dose d' engrais.

Passons maintenant à un problème un petit peu plus complexe et plus intéressant.

On veut estimer les paramètres d'un modèle qui simule l'azote absorbé du colza dans une parcelle agricole en fonction de la dose d'engrais N appliquée sur cette parcelle. On veut que le modèle soit utilisable pour une région.

Ce modèle inclut une variable d'entrée x et deux paramètres. On suppose que lors de l'utilisation du modèle, x sera connu mais ni theta1 ni theta2.

Estimation des paramètres des modèles

Quels paramètres doit-on estimer ?

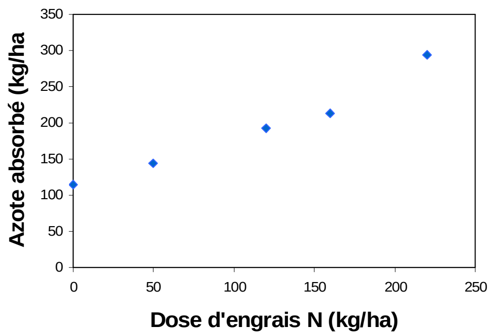
Les deux paramètres du modèle: θ_1 et θ_2

Je reprends ici ma série de questions.

Estimation des paramètres des modèles

Quelle information utiliser ?

Un ***échantillon*** de cinq mesures « d'azote absorbé » obtenues sur cinq parcelles de colza de la ***population*** d'intérêt (une région)



Estimation des paramètres des modèles

Quelle méthode d'estimation utiliser ?

La méthode des ***moindres carrés ordinaires***

Les estimateurs des paramètres sont les valeurs de θ_1 *et* θ_2 qui minimisent

$$\sum_{i=1}^{N}\left(y_i-\theta_1-\theta_2 x_i\right)^2$$

C'est à dire

$$\hat{\theta}_2=\frac{\sum_{i=1}^{N}\left(y_i-\bar{Y}.\right)\left(x_i-\bar{X}.\right)}{\sum_{i=1}^{N}\left(x_i-\bar{X}.\right)^2} \qquad \hat{\theta}_1=\bar{Y}.-\hat{\theta}_2\,\bar{X}.$$

L'estimation des paramètres de ce modèle est un peu plus compliquée que pour le premier problème.

Une méthode assez simple est la méthode des moindres carrés ordinaires.

Avec cette méthode, le estimateurs…

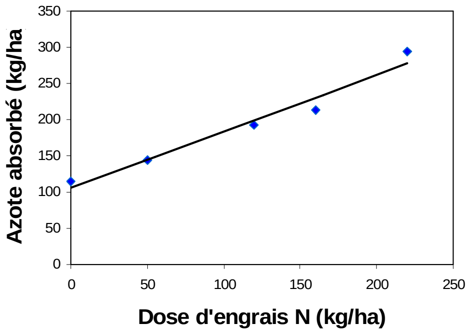
Pour ce modèle particulier, on sait que les valeurs de theta1 et theta2 qui minimisent cette fonction sont définies par les fonctions suivantes…

Peu importe l'expression de ces fonctions. L'idée a retenir est qu'on dispose d'un moyen simple pour calculer les estimateurs des deux paramètres à partir des données.

Ce sera le cas pour tous les modèles linéaires, c'est à dire pour tous les modèles qui correspondent à une combinaison linéaire des paramètres.

Estimation des paramètres des modèles

Ici, avec nos 5 mesures, on obtient $\hat{\theta}_1 = 106.01\ \mathrm{kg.ha^{-1}}$ **et** $\hat{\theta}_2 = 0.78\ \mathrm{kg.kg^{-1}}$



Ces estimateurs sont-ils précis ?

$$E\left[\left(\hat{\theta}-\theta\right)^2\right]=\left[E\left(\hat{\theta}\right)-\theta\right]^2+\operatorname{var}\left(\hat{\theta}\right)$$

Erreur quadratique moyenne — **Biais²** — **Variance**

Ces estimateurs sont-ils précis ?

a. Aspect théorique

« Sous certaines conditions, nos estimateurs sont ***sans biais*** et de ***variances minimales*** parmi les estimateurs sans biais ».

Il faut notamment :

- ***indépendance*** des résidus,

- ***homogénéité*** des variances des résidus.

Ces estimateurs sont-ils précis ?

b. Variances des estimateurs

On peut estimer $\mathrm{var}\left(\hat{\theta}\right)$ à partir des données.

$$\sqrt{\mathrm{var}\left(\hat{\theta}_1\right)} = 11.99 \ \ \mathrm{kg.ha^{-1}}$$

$$\sqrt{\mathrm{var}\left(\hat{\theta}_2\right)} = 0.09 \ \ \mathrm{kg.kg^{-1}}$$

Ces estimateurs sont-ils précis ?

c. Analyse des résidus

$$r_i = y_i - \left(\hat{\theta}_1 + \hat{\theta}_2 x_i\right), \quad i = 1, \ldots, 5$$

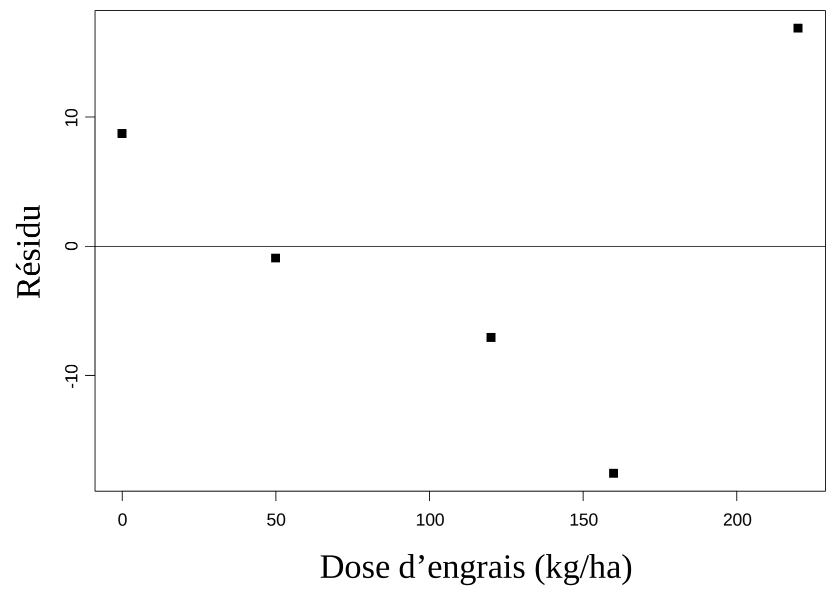
Utile pour vérifier l'indépendance des résidus et l'homogénéité de leurs variances.

Une dernière méthode pour juger de la qualité de la procédure d'estimation est de faire une analyse des résidus.

Un résidu est un écart entre une observation et la valeur correspondante prédite par le modèle. Ici on peut calculer 5 résidus car on dispose de 5 mesures.

L'analyse des résidus est utile pour vérifier qu'il y a bien indépendance des résidus et homogénéité des variances. Si ce n'est pas le cas, alors la méthode des moindres carrés ordinaire n'est pas la méthode conduisant à des estimateurs de variances minimales. Nous verrons plus loin que d'autres méthodes sont plus appropriées.

Estimation des paramètres des modèles



Programme S+

```
DOSE<-c(0,50,120,160,220)
NABS<-c(114.75,144.0,192.38,213,294.16)
DATA<-data.frame(DOSE,NABS)

Fit<-lm(NABS~DOSE,data=DATA)

print(summary(Fit))
plot(DOSE,Fit$residuals,ylab="Residu",ylab="Dose",pch=15)
abline(0,0)
```

Commentaires sur les problèmes 1 et 2

On procède en plusieurs étapes

1. **Quels paramètres estimer ?**
2. **Quelle information disponible ?**
3. **Quelle méthode d’estimation ?**
4. **Quelle est la précision des estimateurs ?**

Commentaires sur les problèmes 1 et 2

C'est facile car

- Modèles linéaires: $\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_p x_p$

  → Relation analytique entre estimateurs et données connue.

- Nombre de données > Nombre de paramètres

- Un seul type de mesure

- Pas de prise en compte d'information *a priori*.

- On a des logiciels pour faire tout ça (SAS, S+, MatLab, ModelMaker…).

En pratique, c'est souvent plus compliqué

- Modèles non linéaires: $\neq \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_p x_p$

  → Relation analytique entre estimateurs et données **inconnue**.

- Peu de données par rapport au nombre de paramètres

- Structure des données complexes

  → plusieurs types de mesures, mesures corrélées

- Information *a priori* parfois disponible.
- Utilisation des logiciels statistiques plus délicate.

Un problème beaucoup plus complexe

→ Modèle non linéaire.

→ Beaucoup de paramètres.

→ Information *a priori*.

→ Différents types de mesures obtenues sur plusieurs parcelles.

Je considère maintenant un problème bcp plus complexe. Habituellement, quand j'ai un peu plus de temps, je présente un problème d'un niveau de complexité intermédiaire mais là je n'ai pas le temps, donc j'attaque directement le cas le plus délicat.

Ce problème est caractérisé par…

Problème 3

Estimer les paramètres du module « fonctionnement potentiel » du modèle AZODYN *(Jeuffroy et Recous, 1999)*

Variables simulées entre sortie hiver et floraison (pas de temps = jour):

- Matière sèche des parties aériennes du blé (kg/ha),
- azote absorbé (kg/ha),
- LAI.

Variables d'entrée:

- Rayonnement global journalier,
- température moyenne journalière de l'air,
- MS et azote absorbé sortie hiver

Estimation des paramètres des modèles

Problème 3

Quelques équations

$$MS_j = MS_{j-1} + \left(E_{b\max} \times ft_{j-1} \times Ei_{j-1} \times C \times RG_{j-1}\right)$$

$$Ei_{j-1} = E_{i\max}\left[1 - \exp\left(-K \times LAI_{j-1}\right)\right]$$

$$LAI_{j-1} = D \times QNc_{j-1}$$

$$MS_j = MS_{j-1} + \left\{E_{b\max} \times C \times E_{i\max}\left[1 - \exp\left(-K \times D \times QNc_{j-1}\right)\right] \times ft_{j-1} \times RG_{j-1}\right\}$$

Les paramètres sont en bleu.

18 paramètres

Paramètre	Signification	Valeur initiale	Gamme
Ebmax	Efficience de conversion du rayonnement	3.3 g/MJ	1.8-4
K	Coefficient d'extinction du rayonnement	0.72	0.6-0.8
D	Rapport LAI / N absorbé critique	0.028	0.02-0.045
Vmax	Vitesse maximale d'absorption d'N	0.5 kg/ha/dj	0.2-0.7
C	PAR/RG	0.48	
Tmin	Température minimale pour photosynthèse	0 °C	
Topt	Température optimale pour photosynthèse	15 °C	
Tmax	Température maximale pour photosynthèse	40 °C	
Eimax	Efficience d'interception du rayonnement	0.96	
Tep-flo	Durée entre épiaison et floraison	150 dj	
E	Paramètre de la courbe critique	1.55 t/ha	
F	Paramètre de la courbe critique	4.4 %	
G	Paramètre de la courbe critique	5.35 %	
H	Paramètre de la courbe critique	-0.442	
L	Paramètre de la courbe max	2 t/ha	
M	Paramètre de la courbe max	6 %	
N	Paramètre de la courbe max	8.3 %	
P	Paramètre de la courbe max	-0.44	

Estimation des paramètres des modèles

Les deux formes d'un modèle dynamique

Forme 1: Système dynamique

$$MS_t = MS_{t-1} + g\left(X_{t-1};\theta\right)$$

Forme 2 : Modèle de réponse

$$MS_t = f\left(t, X;\theta\right)$$

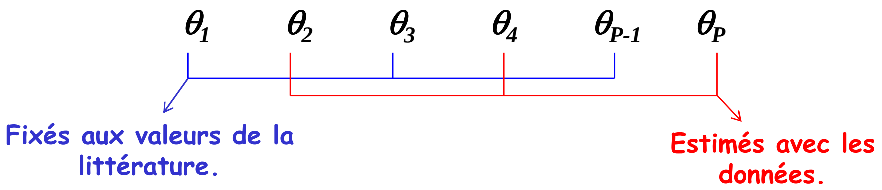
Il est utile à ce stade de rappeler qu'un modèle dynamique peut se mettre sous deux formes…

On vient de voir la première forme. Mais c'est la deuxième forme que l'on va considérer ici, c'est à dire la forme intégrée c'est à dire le modèle de réponse.

Estimation des paramètres des modèles

Quels paramètres estimer ?

Il est nécessaire de sélectionner les paramètres à estimer



- **Problèmes numériques si on estime tous les paramètres.**
- **Même si on pouvait estimer tous les paramètres …**
  … il ne faudrait pas le faire.

Estimer beaucoup de paramètre → variances élevées des estimateurs.

→ augmentation des erreurs de prédictions.

Un problème nouveau que l'on n'a pas rencontrer dans les problèmes précédents est qu'on ne peut pas estimer tous les paramètres du modèle. Il est donc nécessaire de réaliser une sélection.

Nouveau problème: Comment sélectionner les paramètres ?

i. **Utilisation de la littérature.**
ii. **Analyse des équations du modèle.**
iii. **Analyse de sensibilité.**
iv. **Sélection à l'aide de données.**

i. Utilisation de la littérature

« Identifier les paramètres dont les valeurs sont mal connues à partir de la littérature ».

Inconvénients :

- Approche assez subjective.

- Pas toujours de concordance entre les situations considérées dans les articles et celles qui intéressent l'utilisateur.

Estimation des paramètres des modèles

ii. Analyse des équations

« Identifier les paramètres qui ne peuvent pas être estimés simultanément ».

$$MS_j = MS_{j-1} + \left(E_{b\max} \times ft_{j-1} \times Ei_{j-1} \times C \times RG_{j-1} \right)$$

$$Ei_{j-1} = E_{i\max} \left[1 - \exp\left(-K \times LAI_{j-1} \right) \right]$$

$$LAI_{j-1} = D \times QNc_{j-1}$$

$$MS_j = MS_{j-1} + \left\{ E_{b\max} \times C \times E_{i\max} \left[1 - \exp\left(-K \times D \times QNc_{j-1} \right) \right] \times ft_{j-1} \times RG_{j-1} \right\}$$

Quels paramètres si on a uniquement des mesures de MS ?

Quels paramètres si on a uniquement des mesures de MS et de LAI ?

$$MS_j = MS_{j-1} + \left(E_{b\max} \times ft_{j-1} \times Ei_{j-1} \times C \times RG_{j-1} \right)$$

$$Ei_{j-1} = E_{i\max} \left[1 - \exp\left(-K \times LAI_{j-1} \right) \right]$$

$$LAI_{j-1} = D \times QNc_{j-1}$$

$$MS_j = MS_{j-1} + \left\{ E_{b\max} \times C \times E_{i\max} \left[1 - \exp\left(-K \times D \times QNc_{j-1} \right) \right] \times ft_{j-1} \times RG_{j-1} \right\}$$

Uniquement des mesures de MS :

Impossible d'estimer simultanément
- les 3 paramètres $E_{b\max}, C, E_{i\max}$
- les 2 paramètres K, D

Mesures de MS et de LAI :

Impossible d'estimer simultanément les 3 paramètres $E_{b\max}, C, E_{i\max}$

iii. Analyse de sensibilité

« Sélectionner les paramètres qui ont une forte influence sur les variables simulées par le modèle ».

Inconvénients :

Il faut définir un seuil de sensibilité.

Ne permet pas de diagnostiquer les problèmes d'identifiabilité.

Estimation des paramètres des modèles

iv. Utilisation des données

« Sélectionner les paramètres qu'il faut estimer pour optimiser la qualité prédictive du modèle (Wallach et al., 2001) ».

Nb de paramètres estimés avec des données	MSEP*vc*
1	$MSEP_1$
2	$MSEP_2$
3	$MSEP_3$
…	…
P	$MSEP_P$

Sélection des paramètres qu'il faut estimer pour minimiser le MSEP

Quels paramètres estimer ?

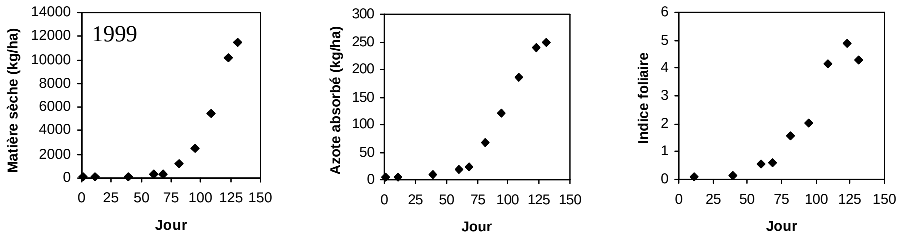
- **13 paramètres** sont fixés aux valeurs fournies par la littérature.
- **Un paramètre** est fixé après analyse des équations.
- **Quatre paramètres** sont estimés à partir des données : E_{BMAX}, D, K et V_{MAX}

Quelle information disponible ?

• Mesures de **matière sèches** des parties aériennes du blé, de **LAI** et **d'azote absorbé** obtenues à Grignon pour 6 années.

• Dix dates de mesure chaque année entre sortie-hiver et floraison.

• Trois répétitions à chaque date. On utilise les moyennes.



Quelle méthode d'estimation utiliser ?

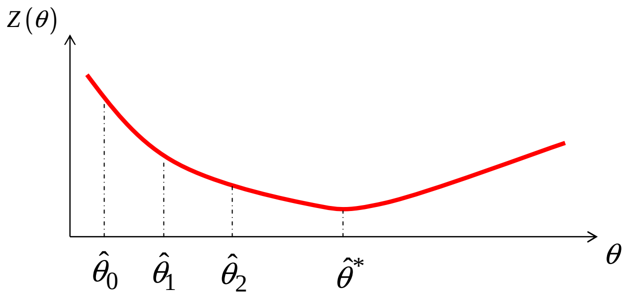
1er possibilité : La méthode des moindres carrés ordinaires

Trouver la valeur de θ qui minimise : $Z(\theta) = \sum_{i=1}^{N} \left[y_i - f(t_i, x_i; \theta) \right]^2$

Problème :

- le modèle est non linéaire,
- on ne peut pas trouver l'expression analytique des estimateurs.

Appliquer la méthode des MCO avec un algorithme itératif

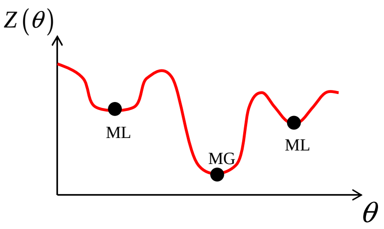


La solution consiste à appliquer la méthode à l'aide d'un algorithme itératif. Ce graphique présente en abscisse la valeur d'un paramètre hypothétique theta et en ordonnée la valeur du critère des moindres carrés en fonction de theta. La courbe présente un minimum et on cherche à déterminer la valeur du paramètre qui correspond à ce minimum.

Pour cela, on part d'une valeur initiale theta0 que l'utilisateur doit fournir. Puis on va s'approcher de la valeur optimale theta* en passant par des valeurs intermédiaires theta1, theta2…

Estimation des paramètres des modèles

Minimum locaux et minimum globaux



→ Essayez plusieurs valeurs initiales !

En pratique, la fonction Z est souvent moins idéale. Elle présente souvent des minimums locaux que l'on va chercher à éviter et un minimum global qu'on va chercher à atteindre. Pour éviter de tomber dans un minimum local, il est nécessaire de faire tourner l'algorithme d'estimation plusieurs fois à partir de plusieurs valeurs initiales différentes.

Aspect pratique

• On peut utiliser un logiciel statistique (SAS, S+, MatLab, bibliothèques Fortran ou C++…)

• On donne :

- des données,
- un modèle,
- des valeurs initiales des paramètres.

• Le logiciel fournit en sortie les valeurs estimées des paramètres.

Quelle méthode d'estimation utiliser ?

1er possibilité : La méthode des moindres carrés ordinaires

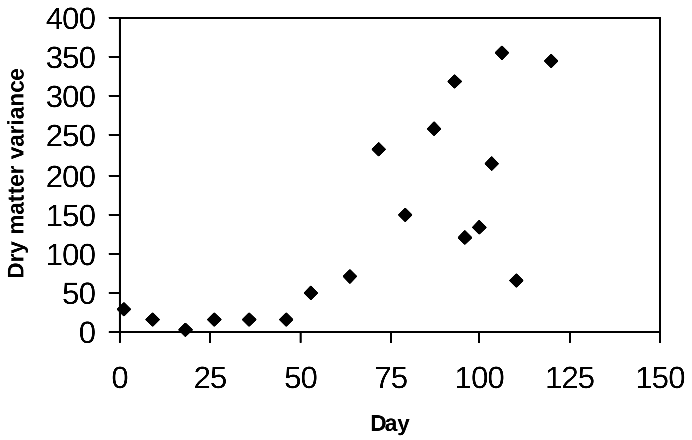
Trouver la valeur de θ qui minimise : $Z(\theta) = \sum_{i=1}^{N} \left[y_i - f(t_i, x_i; \theta) \right]^2$

Inconvénient :

- Les estimateurs ne sont pas de variances minimales si les résidus ont des variances hétérogènes.

- Or, ici, il y a plusieurs types de mesures et la variance des mesures dépend de la date d'observation.

Les variances des mesures sont hétérogènes



Quelle méthode d'estimation utiliser ?

La méthode des moindres carrés pondérés

Trouver la valeur de θ qui minimise :

$$Z(\theta) = \sum_{i=1}^{N} \frac{\left[y_i - f(t_i, x_i; \theta)\right]^2}{\sigma_i^2}$$

$$\text{avec } \hat{\sigma}_i^2 = \frac{1}{K(K-1)} \sum_{k=1}^{K} (y_{ik} - y_i)^2$$

La seule différence est qu'on divise chaque terme par la variance des observations.

L'idée est de donner moins de poids aux mesures dont la variance est grande.

Estimation des paramètres des modèles

Quelle méthode d'estimation utiliser ?

La méthode des moindres carrés pondérés

On minimise

$$Z_{MCP}(\theta)=\sum_{i=1}^{6}\sum_{j=1}^{10}\frac{\left[y_{ij}^{MS}-f^{MS}\left(t_j,x_i;\theta\right)\right]^2}{\hat{\sigma}_{MS.ij}^2}+\sum_{i=1}^{6}\sum_{j=1}^{10}\frac{\left[y_{ij}^{N}-f^{N}\left(t_j,x_i;\theta\right)\right]^2}{\hat{\sigma}_{N.ij}^2}+\sum_{i=1}^{6}\sum_{j=1}^{10}\frac{\left[y_{ij}^{L}-f^{L}\left(t_j,x_i;\theta\right)\right]^2}{\hat{\sigma}_{L.ij}^2}$$

$$\hat{\sigma}_{MS.ij}^2=\frac{1}{K\left(K-1\right)}\sum_{k=1}^{K}\left[y_{ijk}^{MS}-y_{ij}^{MS}\right]^2$$

La méthode des MCP est appliquée à notre modèle dynamique. L'indice i est l'indice de l'année de mesure et l'indice j est l'indice de la date de mesure.

Ici on a trois types de mesures yijMS, …

On calcule la différence entre chaque mesure et la prédiction du modèle correspondante.

On élève la différence au carré et on la divise par la variance de la mesure.

Chaque mesure correspond à la moyenne de trois répétitions. On peut calculer les variances à partir des valeurs élémentaires obtenues sur les répétitions.

Estimation des paramètres des modèles

Application des moindres carrés pondérés pour estimer les 4 paramètres

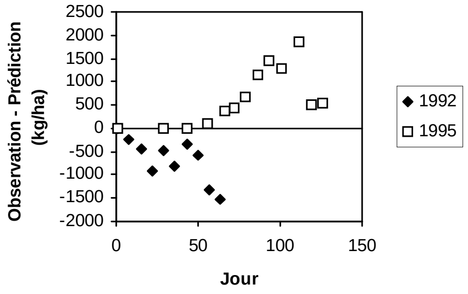
Paramètre	Valeur initiale	Valeur estimée MCP
E_{BMAX} (g/MJ)	3.3	3.29 (0.11)
D	0.028	0.037 (0.06)
K	0.72	0.74 (0.001)
V_{MAX} (kg/ha/dj)	0.5	0.38 (0.02)

Estimation des paramètres des modèles

Ces estimateurs sont-ils précis ?

Analyse des résidus obtenus avec la méthodes des moindres carrés pondérés

Les résidus ne sont pas indépendants



Méthodes pour prendre en compte les corrélations

- Moindres carrés généralisés.
- Modèles mixtes.

Un dernier problème

Comment prendre en compte l'information *a priori* ?

• Jusqu'à présent, nos quatre paramètres sont estimés à partir des données, sans utiliser l'information a priori.

• Les ***méthodes Bayésiennes*** sont utiles pour estimer les paramètres à la fois à partir des données et de l'information a priori.

Conclusion

On procède en plusieurs étapes

1. Quels paramètres estimer ?

→ Dans les cas simples, on peut tout estimer.

→ Dans les cas complexes, il faut faire une sélection.

2. Quelle information disponible ?

→ Les données

→ Information a priori

3. Quelle méthode d’estimation ?

→ Moindres carrés ordinaires,

→ Moindres carrés pondérés/généralisés,

→ Méthodes Bayésiennes…

4. Quelle est la précision des estimateurs ?

→ Aspects théoriques, variances, résidus.