Plans avancés

- Plans pour variables qualitatives à plus de 2 modalités
- Plans pour variables quantitatives et qualitatives
- Plans optimaux

François Husson

Unité pédagogique de mathématiques appliquées
Agrocampus Ouest

Plans symétriques

- Plans symétriques : tous les facteurs ont le même nombre de modalités : carrés latin, carré gréco-latins, MOLS
- Plans asymétriques : le nombre de modalités n'est pas le même pour tous les facteurs

Carrés latins

Objectif : étudier 3 facteurs à J modalités

Mode de construction :

- 1ère colonne, numéroter de 0 à $J-1$
- Colonnes suivantes, ajouter 1 à la colonne précédente, modulo J

Plan à 3 facteurs à 3 modalités
Plan complet $3^3 = 27$ essais

F1 \ F2	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

F1	F2	F3
0	0	0
0	1	1
0	2	2
1	0	1
1	1	2
1	2	0
2	0	2
2	1	0
2	2	1

Plan à 3 facteurs à 5 modalités
Plan complet $5^3 = 125$ essais

F1 \ F2	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

F1	F2	F3
0	0	0
0	1	1
0	2	2
0	3	3
0	4	4
1	0	1
1	0	2
...	...	...
4	4	3

Propriétés d'un carré latin

- Effets principaux orthogonaux (vérification par tableau croisé)

F1 \ F2	0	1	2
0	1	1	1
1	1	1	1
2	1	1	1

F1 \ F3	0	1	2
0	1	1	1
1	1	1	1
2	1	1	1

F2 \ F3	0	1	2
0	1	1	1
1	1	1	1
2	1	1	1

- Effets principaux orthogonaux avec leurs interactions mais partiellement confondus avec interactions qui ne les concernent pas : $F_1 \perp F_1F_2$ $\quad F_1 \perp F_1F_3$ mais on n'a pas $F_1 \perp F_2F_3$
- Etude du plan par analyse de variance à 3 facteurs
- Grande confiance dans l'estimation des paramètres, pas dans les tests (peu de ddl)
- Si optimum potentiellement atteint pour une expérience non testée $\Rightarrow$ la tester et si valeur prédite proche de valeur observée, le modèle est validé

Carrés latins mutuellement orthogonaux (MOLS)

- Juxtaposition de MOLS pour étudier plus de facteurs
- J premier ou puissance d'un nb premier $\Rightarrow$ il existe $J-1$ MOLS

Mode de construction de carrés latins orthogonaux :

- si J est un nombre premier :
  - construire un premier carré latin
  - pour construire un autre carré latin, prendre $\alpha \in \{2, ..., J-1\}$ et mettre ligne i colonne j la valeur $\alpha \times i + j$ modulo J.
    Construire $J-1$ carrés avec les valeurs de α entre 2 et $J-1$

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\alpha = 2$

	0	1	2
0	0	1	2
1	2	0	1
2	1	2	0

$2 \times 0 + 1 = 1$
$2 \times 1 + 2 = 1$
$2 \times 2 + 1 = 2$

- si J est puissance d'un nombre premier : utilisation de tables

Carrés latins mutuellement orthogonaux : exemple $J = 3$

Exemple avec 3 modalités :

	0	1	2
0	a	b	c
1	b	c	a
2	c	a	b

	0	1	2
0	α	β	γ
1	γ	α	β
2	β	γ	α

$\Rightarrow$

	0	1	2
0	aα	bβ	cγ
1	bγ	cα	aβ
2	cβ	aγ	bα

$$\begin{bmatrix} 0 & 0 & a & \alpha \\ 0 & 1 & b & \beta \\ 0 & 2 & c & \gamma \\ 1 & 0 & b & \gamma \\ \vdots & \vdots & \vdots & \vdots \\ 2 & 1 & a & \gamma \\ 2 & 2 & b & \alpha \end{bmatrix}$$

La juxtaposition de 2 carrés latins orthogonaux conduit à un carré gréco-latin permettant d'étudier 4 facteurs

Carrés latins mutuellement orthogonaux : exemple $J = 5$

	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\alpha = 2$

	0	1	2	3	4
0	0	1	2	3	4
1	2	3	4	0	1
2	4	0	1	2	3
3	1	2	3	4	0
4	3	4	0	1	2

$\alpha = 3$

	0	1	2	3	4
0	0	1	2	3	4
1	3	4	0	1	2
2	1	2	3	4	0
3	4	0	1	2	3
4	2	3	4	0	1

$\alpha = 4$

	0	1	2	3	4
0	0	1	2	3	4
1	4	0	1	2	3
2	3	4	0	1	2
3	2	3	4	0	1
4	1	2	3	4	0

Possibilité d'étudier jusqu'à 6 facteurs à 5 modalités

Carrés latins mutuellement orthogonaux : tables

d'ordre 4

0	1	2	3
1	2	3	0
2	3	0	1
3	0	1	2

0	1	2	3
2	3	0	1
3	2	1	0
1	0	3	2

0	1	2	3
3	2	1	0
1	0	3	2
2	3	0	1

Pas de carrés latin mutuellement orthogonaux d'ordre 6

d'ordre 8

0	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	0
3	4	5	6	7	0	1
4	5	6	7	0	1	2
5	6	7	0	1	2	3
6	7	0	1	2	3	4
7	0	1	2	3	4	5

0	1	2	3	4	5	6
2	3	4	5	6	7	0
4	5	6	7	0	1	2
6	7	0	1	2	3	4
3	4	5	6	7	0	1
1	2	3	4	5	6	7
7	0	1	2	3	4	5
5	6	7	0	1	2	3

0	1	2	3	4	5	6
3	4	5	6	7	0	1
6	7	0	1	2	3	4
5	6	7	0	1	2	3
7	0	1	2	3	4	5
4	5	6	7	0	1	2
1	2	3	4	5	6	7
2	3	4	5	6	7	0

0	1	2	3	4	5	6
4	5	6	7	0	1	2
3	4	5	6	7	0	1
7	0	1	2	3	4	5
6	7	0	1	2	3	4
2	3	4	5	6	7	0
5	6	7	0	1	2	3
1	2	3	4	5	6	7

0	1	2	3	4	5	6
5	6	7	0	1	2	3
1	2	3	4	5	6	7
4	5	6	7	0	1	2
2	3	4	5	6	7	0
7	0	1	2	3	4	5
3	4	5	6	7	0	1
6	7	0	1	2	3	4

0	1	2	3	4	5	6
6	7	0	1	2	3	4
7	0	1	2	3	4	5
1	2	3	4	5	6	7
5	6	7	0	1	2	3
3	4	5	6	7	0	1
2	3	4	5	6	7	0
4	5	6	7	0	1	2

0	1	2	3	4	5	6
7	0	1	2	3	4	5
5	6	7	0	1	2	3
2	3	4	5	6	7	0
1	2	3	4	5	6	7
6	7	0	1	2	3	4
4	5	6	7	0	1	2
3	4	5	6	7	0	1

Carrés latins mutuellement orthogonaux

Nombre de facteurs qu'il est possible d'étudier

Facteurs à 3 modalités

r	Nb essais 3^r	Résolution III	IV	V
2	9	4		
3	27	13	4	
4	81	40	10	5
5	243	121	?	11

Facteurs à 4 modalités

Nb essais 4^r	Résolution III	IV
16	5	
64	21	6

Exemple d'utilisation des MOLS en analyse sensorielle

Sélectionner les colonnes dans un MOLS pour construire des plans en bloc incomplets équilibrés.
Le numéro dans une case correspond au numéro de produit dégusté.
Le facteur produit est orthogonal au facteur juge et au facteur rang.

Juges \ Rang	1	2	3	4	5
1	1	2	5	3	4
2	2	3	1	4	5
3	3	4	2	5	1
4	4	5	3	1	2
5	5	1	4	2	3

Juges \ Rang	1	2	3	4	5
11	1	3	4	5	2
12	2	4	5	1	3
13	3	5	1	2	4
14	4	1	2	3	5
15	5	2	3	4	1

Juges \ Rang	1	2	3	4	5
6	1	5	2	4	3
7	2	1	3	5	4
8	3	2	4	1	5
9	4	3	5	2	1
10	5	4	1	3	2

Juges \ Rang	1	2	3	4	5
16	1	4	3	2	5
17	2	5	4	3	1
18	3	1	5	4	2
19	4	2	1	5	3
20	5	3	2	1	4

Plans avec des facteurs à 2 et 4 niveaux

- Voir chaque facteur à 4 niveaux comme 2 facteurs à 2 niveaux
- Construire un plan fractionnaire 2^{p-k}
- Utiliser 2 facteurs à 2 niv pour coder un facteur à 4 niv :

$$\begin{array}{rr} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{array} \quad \Longrightarrow \quad \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}$$

En $N = 4^r$ essais, on peut étudier n_4 facteurs à 4 modalités avec $0 \leq n_4 \leq (4^r - 1)/3$ et $(N - (3 \times n_4 + 1))$ facteurs à 2 modalités
Ex : en 16 essais, $n_4 = 5$ et $n_2 = 0$ ou $n_4 = 3$ et $n_2 = 6$

Avec des facteurs à 8 niveaux, on utilise 3 facteurs à 2 niveaux

Plans avec des facteurs à 2, 3 et 4 niveaux

- Considérer les facteurs à 3 niveaux comme des facteurs à 4 niveaux
- Construire le plan d'expériences avec des facteurs à 2 et 4 niveaux (voir diapo précédente)
- Pour chaque facteur à 3 niveaux, prendre un facteur à 4 niveaux et compresser 2 niveaux

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \quad \Longrightarrow \quad \begin{array}{c} 1 \\ 2 \\ 2 \\ 3 \end{array}$$

$\Longrightarrow$ Vérifier la qualité d'un plan en calculant $(X'X)^{-1}$

Plans avec des nombres de modalités différents

Exemple : 2 facteurs à 2 modalités, 1 facteur à 3 et 1 facteur à 5

- nb max d'essais : $2^2 * 3 * 5 = 60$ pour le plan complet
- nb min d'essais : nb de paramètres à estimer :
  $1 + 2 \times (2-1) + (3-1) + (5-1) = 9$
- Pour l'orthogonalité des facteurs, il faut
  - $F_1 \perp F_2$ donc minimum $2 \times 2 = 4$ essais
  - $F_1 \perp F_3$ et $F_2 \perp F_3$ donc minimum $2 \times 3 = 6$ essais
  - $F_1 \perp F_4$ et $F_2 \perp F_4$ donc minimum $2 \times 5 = 10$ essais
  - $F_3 \perp F_4$ donc minimum $3 \times 5 = 15$ essais

  $\Rightarrow \quad N \geq PPCM(4, 6, 10, 15) = 60 \quad \Rightarrow$ seul le plan complet permet d'estimer sans confusion d'effets

Plans avec des nombres de modalités différents : $L_{36}2^33^3$

Plan $L_{36}2^33^3 \Rightarrow N_{max} = 216,\ N_{min} = 10$ et $N \geq PPCM(4,9) = 36$

Mode de construction : construire un plan 3^{3-1} et un plan 2^{3-1} puis introduire le plan 2^{3-1} pour chaque ligne du plan 3^{3-1}

Plan 3^{3-1}

0	**0**	**0**
1	0	1
2	0	2
0	1	1
1	1	2
2	1	0
0	2	2
1	2	0
2	2	1

Plan 2^{3-1}

0	0	0
0	1	1
1	0	1
1	1	0

	F1	F2	F3	F4	F5	F6
exp1	**0**	**0**	0	0	0	0
exp2	**0**	**0**	0	0	1	1
exp3	**0**	**0**	0	1	0	1
exp4	**0**	**0**	0	1	1	0
exp5	1	0	1	0	0	0
exp6	1	0	1	0	1	1
exp7	1	0	1	1	0	1
exp8	1	0	1	1	1	0
exp9	2	0	2	0	0	0
exp10	2	0	2	0	1	1
exp11	2	0	2	1	0	1
exp12	2	0	2	1	1	0
exp13	0	1	1	0	0	0
exp14	0	1	1	0	1	1
exp15	0	1	1	1	0	1
exp16	0	1	1	1	1	0
exp17	1	1	2	0	0	0
exp18	1	1	2	0	1	1
exp19	1	1	2	1	0	1
exp20	1	1	2	1	1	0

(suite)

exp21	2	1	0	0	0	0
exp22	2	1	0	0	1	1
exp23	2	1	0	1	0	1
exp24	2	1	0	1	1	0
exp25	0	2	2	0	0	0
exp26	0	2	2	0	1	1
exp27	0	2	2	1	0	1
exp28	0	2	2	1	1	0
exp29	1	2	0	0	0	0
exp30	1	2	0	0	1	1
exp31	1	2	0	1	0	1
exp32	1	2	0	1	1	0
exp33	2	2	1	0	0	0
exp34	2	2	1	0	1	1
exp35	2	2	1	1	0	1
exp36	2	2	1	1	1	0

Prise en compte d'un effet bloc

Exemple d'effet bloc : effet animal, effet parcelle

Pour construire un plan, un effet bloc est un effet comme un autre $\Rightarrow$ l'introduire dans les facteurs et le prendre en compte pour construire le plan (comme d'habitude)

Pour l'analyse des résultats, considérer l'effet bloc dans le modèle (mais pas d'interactions avec l'effet bloc)

Construction de plans avec R

```
# Construction de plans fractionnaires à 2 modalités
library(FrF2)
plan1 <- FrF2(nruns=8, nfactors=4, factor.names=list(temp=c("min","max"),
     press=c("low","normal"),material=c("M1","M2"),state=c("new","aged")))
plan2 <- FrF2(nfactors=5, resolution=5)
summary(plan2)

# Construction de plans orthogonaux
library(DoE.base)
fac.design(nlevels=c(4,3,3,2))  # nb d'essais calculé pour plan fractionnaire
fac.design(nlevels=c(2,2,3,3,6), blocks=6, seed=12345)
oa.design(nlevels=c(2,2,2,3,3,3), nruns=36) # plan orthogonal

## Vérification de la qualité d'un plan selon le modèle voulu
options(contrasts=c("contr.sum","contr.sum"))
X <- model.matrix(~A+B+C+D+E+A:B+A:C+B:C, data=plan2)
solve(t(X)%*%X)
```

Construction de plans avancés

```
library(DoE.base)
# Comparer 6 variétés (A,B,C,D,E,F) en utilisant 4 blocs de 6 parcelles
plan <- oa.design(nlevels=c(6,4),factor.names=list(variete=LETTERS[1:6],
     bloc=1:4))

# Comparer 3 variétés (A,B,C), 2 doses d'engrais (1,2) en 6 lignes * 6 col
plan <- oa.design(nlevels=c(6,6,3,2))

library(planor)
# Analyse de sensibilité d'un modèle d'épidémiologie animale
# Plan d'expérience numérique : 12 facteurs à 4 niveaux, 7 facteurs à 2 niv
# Modèle : effets principaux + interactions entre 2 facteurs
# 4^12*2^7=2^31 combinaisons possibles, moins de 2^12 = 4096 expé
frac.key <- planor.designkey(factors = LETTERS[1:19],
     nlevels = c(rep(4,12),rep(2,7)),
     model = ~(A+B+C+D+E+F+G+H+I+J+K+L+M+N+O+P+Q+R+S)^2, nunits = 4096)
frac.plan <- planor.design(frac.key)
```


Mélange de variables quantitatives et qualitatives

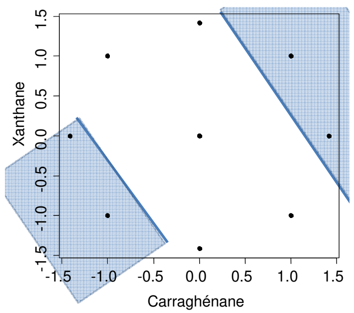
Deux stratégies possibles :

- transformer les variables quanti en quali avec au moins 3 niveaux puis construire le plan pour variables quali
- on construit un plan avec les variables quanti seules et un plan avec les variables quali seules puis on répète le plan quanti pour chaque expérience du plan quali $\Rightarrow$ nb d'essais très important (stratégie possible si très peu de variables quali)

Dans les 2 cas, on analyse les données avec un modèle de covariance

Plans optimaux

- 1 facteur à 2 modalités, 2 facteurs à 3 et 2 facteurs à 4 :
  $n_{max} = 288$, $PPCM(6, 8, 9, 16, 12) = 144$
  Trade of / analyse conjointe en marketing : 16 essais possibles
  On perd l'orthogonalité mais comment prévoir avec le plus de précision ?
- Y viscosité
  x_1 teneur en carraghénane
  x_2 teneur en xanthane
  Pour certaines recettes la viscosité sera trop forte (ou trop faibles)
  $\Rightarrow$ impossible de faire ces recettes



Plans optimaux : les critères

$$\mathbb{V}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$\Longrightarrow$ Minimiser $(X'X)^{-1}$: problème, il faudrait un seul critère

- $(X'X)$ matrice d'information
- $(X'X)^{-1}$ matrice de dispersion

Trois critères :

- D-optimalité : minimise le déterminant de la matrice de dispersion $\Leftrightarrow$ maximise la matrice d'information
- A-optimalité : minimiser la moyenne de la variance des coefficients de la matrice de dispersion
- G-optimalité : trouver les expériences qui prévoient avec le plus de précision $\Rightarrow$ minimiser la variance de prédiction

Le critère de D-optimalité est plus rapide à calculer

Plans optimaux : algorithme d'échange

Algorithme d'échange de Fedorov :

1. définir un grand nombre d'expériences potentielles N
2. définir le nombre d'essais à réaliser n
3. tirer au hasard n expériences parmi les N
4. calculer le critère choisi : $\det(X'X)$ par exemple
5. sortir au hasard 1 des n expériences du plan et en ajouter 1 des $N - n$ (au hasard)
6. si le déterminant augmente, conserver cet échange, sinon annuler l'échange
7. itérer les étapes 5 et 6 jusqu'à convergence

Plans optimaux : avantages /inconvénients

Avantages :

- plans très flexibles par rapport au nombre d'essais
- on peut imposer certains essais
- on peut rajouter des expériences au cours du plan

Inconvénients :

- fournit toujours un plan, sans garantie sur sa qualité $\Rightarrow$ toujours vérifier la qualité
- la convergence vers l'optimum global n'est pas assurée $\Rightarrow$ relancer plusieurs fois l'algorithme

Construction de plans optimaux avec R

```
# EXEMPLE 1: modèle quadratique avec 3 variables
library(AlgDesign)
dat<-gen.factorial(levels=3,nVars=3,varNames=c("A","B","C"))
desD<-optFederov(~quad(A,B,C),data=dat,nTrials=14,eval=TRUE)

levels<-seq(-1,1,by=.1)
dat<-expand.grid(list(A=levels,B=levels,C=levels)) ## grille avec 9261 essais
desL <- optFederov(~quad(.), data=dat, nTrials=14, eval=TRUE)

# EXEMPLES 2 : plan fractionnaire 2^{4-1}
dat <- gen.factorial(levels=2,nVars=4,varNames=LETTERS[1:4])
desH <- optFederov(~.,data=dat,8)

# Plan orthogonal de Plackett-Burman
dat<-gen.factorial(levels=2,nVars=11,varNames=LETTERS[1:11]) # 2048 essais
desPB<-optFederov(~.,data=dat,12,nRepeats=20)
X <- model.matrix(~.,data=desPB$design) ## pour vérifier l'orthogonalité
t(X)%*%X

# Construction d'un carré latin (il faut nRepeats suffisamment grand)
lv<-factor(1:5)
dat<-expand.grid(A=lv,B=lv,C=lv)
desL<-optFederov(~.,data=dat,nTrials=25,nRepeats=100)
```


Construction de plans optimaux avec R

```
# EXEMPLE 3: essais imposés
dat<-gen.factorial(levels=3,nVars=3,varNames=c("A","B","C"))
desD<-optFederov(~quad(A,B,C),data=dat,nTrials=14,eval=TRUE)

# ajout d'essais au plan précédent
dat<-gen.factorial(levels=3,nVars=3,varNames=c("A","B","C"))
desA<-optFederov(~quad(.), data=dat, nTrials=25, augment=TRUE, rows=desD$rows)

# Le plan desH est complété pour prendre en compte une interaction:
dat <- gen.factorial(levels=2,nVars=5,varNames=LETTERS[1:5])
desH <- optFederov(~., data=dat, nTrials=8)
desH2 <- optFederov(~A+B+C+D+E+I(A*B), data=dat, nTrials=10,
        augment=TRUE, rows=desH$rows)
```