Optimal design of blocked and split-plot experiments for fixed-effects and variance-components estimation

Peter Goos

Kalliopi Mylona

Universiteit Antwerpen

Optimal design of blocked and split-plot **response surface** experiments for fixed-effects and variance-components estimation

Peter Goos

Kalliopi Mylona

Universiteit Antwerpen

Outline

- blocked and split-plot experiments
- model
- optimal design approach
  - fixed effects
  - variance components
- computational aspects
  - Gauss-Stieltjes-Wigert quadrature
  - Gauss-Jacobi quadrature
- computational results

Blocked experiments

- useful whenever experimental runs cannot be performed under homogeneous circumstances
- runs are partitioned in groups or blocks
- runs within a block are more alike than runs from different blocks
- blocks can be days, batches, test subjects, …
- one way of grouping runs (one blocking factor)

Split-plot experiments

- hard-to-change factors
  - levels are reset only a couple of times
  - gives rise to blocks of observations for which the hard-to-change factors are not reset
  - these blocks are called whole plots
- two-stage production processes
  - experiments often span more than one step in a production process
  - some factors are applied in a first stage
  - others are applied in a subsequent stage

Model

$$y_{ij} = \mathbf{f}'(\mathbf{x}_{ij})\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}$$

experimental factors

Model

$$y_{ij} = \mathbf{f}'(\mathbf{x}_{ij})\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}$$

block effect

i.i.d. $N(0, \sigma_\gamma^2)$

random error

i.i.d. $N(0, \sigma_\varepsilon^2)$

Model

$$y_{ij} = \mathbf{f}'(\mathbf{w}_i, \mathbf{s}_{ij})\beta + \gamma_i + \varepsilon_{ij}$$

whole plot factors

sub-plot factors

Model

$$y_{ij} = \mathbf{f}'(\mathbf{w}_i, \mathbf{s}_{ij})\beta + \gamma_i + \varepsilon_{ij}$$

whole-plot effect

i.i.d. $N(0, \sigma_\gamma^2)$

random error

i.i.d. $N(0, \sigma_\varepsilon^2)$

Model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \varepsilon$$

$$\begin{aligned}\mathbf{V} &= \text{var}(\mathbf{Y}) \\ &= \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\gamma^2 \mathbf{Z}\mathbf{Z}'\end{aligned}$$

Model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \varepsilon$$

$$\begin{aligned}\mathbf{V} &= \mathrm{var}(\mathbf{Y}) \\ &= \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\gamma^2 \mathbf{Z}\mathbf{Z}' \\ &= \sigma_\varepsilon^2 (\mathbf{I}_n + \eta \mathbf{Z}\mathbf{Z}') \text{ where } \eta = \frac{\sigma_\gamma^2}{\sigma_\varepsilon^2}\end{aligned}$$

Model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \varepsilon$$

$$\begin{aligned}
\mathbf{V} &= \mathrm{var}\left(\mathbf{Y}\right) \\
&= \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\gamma^2 \mathbf{Z}\mathbf{Z}' \\
&= \sigma_\varepsilon^2 \left(\mathbf{I}_n + \eta \mathbf{Z}\mathbf{Z}'\right) \text{ where } \eta = \frac{\sigma_\gamma^2}{\sigma_\varepsilon^2} \\
&= \sigma_\varepsilon^2 \left(\mathbf{I}_n + \frac{\rho}{1-\rho} \mathbf{Z}\mathbf{Z}'\right) \text{ where } \rho = \frac{\sigma_\gamma^2}{\sigma_\gamma^2 + \sigma_\varepsilon^2}
\end{aligned}$$

Estimation

- generalized least squares regression

$$\hat{\beta} = \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{Y}$$

- covariance matrix

$$\text{var}(\hat{\beta}) = \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X}\right)^{-1}$$

- $\hat{\mathbf{V}}$ is the estimated covariance structure of the observations
- $\hat{\mathbf{V}}$ can be estimated using restricted maximum likelihood estimation (REML)

Design approaches

- split-plot designs:
  - stratum-by-stratum construction: Trinca & Gilmour (2001)
  - equivalent-estimation designs: Vining, Kowalski & Montgomery (2005), Parker, Kowalski & Vining (2006, 2007a, 2007b)
  - optimal design approach
- blocked experiments:
  - orthogonal blocking
  - optimal design approach

Optimal design

- blocked experiments: Goos & Vandebroek (2001)
- split-plot experiments: Goos & Vandebroek (2003), Jones & Goos (2007)
- D-optimality criterion: seeks designs that maximize determinant of information matrix on $\boldsymbol{\beta}$

$$\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

Weakness 1

8 runs

4 blocks of size 2

2 factors

full quadratic model

8 parameters

variance components
are not estimable !

Blocks	x_1	x_2
1	-1	1
	1	0
2	0	1
	-1	-1
3	0	-1
	1	1
4	1	-1
	-1	0

Weakness 2

- design depends on variance components through **V**
- not on their absolute magnitude, but on their relative magnitude
- unelegant, but …
  - in many cases, optimal design does not depend on η or ρ
  - in other cases, a design that is optimal for one value of η or ρ is highly efficient for other values
- setting $\eta = 1$ or $\rho = 0.5$ is recommended

Solutions

When constructing optimal designs

1. take into account information matrix for variance components

2. use a prior distribution for either η or ρ

Variance components

- restricted maximum likelihood estimation

Variance components

- restricted maximum likelihood estimation
- expression for information matrix can be found in, for instance, Searle, Casella & McCulloch (1992):

$$\frac{1}{2}\begin{bmatrix} \mathrm{tr}\left((\mathbf{PZZ'})^2\right) & \mathrm{tr}\left(\mathbf{P}^2\mathbf{ZZ'}\right) \\ \mathrm{tr}\left(\mathbf{P}^2\mathbf{ZZ'}\right) & \mathrm{tr}\left(\mathbf{P}^2\right) \end{bmatrix}$$

Variance components

- restricted maximum likelihood estimation
- expression for information matrix can be found in, for instance, Searle, Casella & McCulloch (1992):

$$\frac{1}{2}\begin{bmatrix} \operatorname{tr}\left((\mathbf{PZZ}')^2\right) & \operatorname{tr}\left(\mathbf{P}^2\mathbf{ZZ}'\right) \\ \operatorname{tr}\left(\mathbf{P}^2\mathbf{ZZ}'\right) & \operatorname{tr}\left(\mathbf{P}^2\right) \end{bmatrix}$$

with

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}$$

Composite criterion

- find design that maximizes

$$D = \left|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right|^{1/p} \times \begin{vmatrix} \mathrm{tr}\left((\mathbf{PZZ}')^2\right) & \mathrm{tr}\left(\mathbf{P}^2\mathbf{ZZ}'\right) \\ \mathrm{tr}\left(\mathbf{P}^2\mathbf{ZZ}'\right) & \mathrm{tr}\left(\mathbf{P}^2\right) \end{vmatrix}^{1/2}$$

with

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}$$

- depends on η or ρ too, so D should read $D(\eta)$ or $D(\rho)$

Bayesian composite criterion

- assume prior distribution for either η or ρ
- find design that maximizes

$$D_B = \int_0^\infty \log(D(\eta))\,\pi(\eta)\,d\eta$$

or

$$D_B = \int_0^1 \log(D(\rho))\,\pi(\rho)\,d\rho$$

with $\pi(\eta)$ and $\pi(\rho)$ the prior distributions

Bayesian **composite criterion**

- no analytical expressions for these integrals
- need to be evaluated numerically
- usually done using pseudo-Monte Carlo sampling
- for instance

$$D_B = \sum_{i=1}^{R} \log\left(D\left(\eta_i\right)\right)$$

where η_i is a random draw from the prior distribution $\pi(\eta)$ and R is large

Lognormal prior for η

- Bayesian composite criterion

$$D_B = \int_0^{\infty} \log\left(D(\eta)\right) \frac{1}{\eta\sigma\sqrt{2\pi}} e^{-\{\log(\eta)-\mu\}^2/2\sigma^2} d\eta$$

- Gauss-Hermite quadrature
- Gaussian quadrature based on Stieltjes-Wigert polynomials

$$D_B \approx \sum_{i=1}^{R} w_i^{GSW} \log\left(D\left(e^{\mu}\alpha_i^{GSW}\right)\right)$$

- R is small

Choice of μ and σ^2

- geometric mean: $m = e^{\mu}$
- geometric standard deviation: $s = e^{\sigma}$
- 99.7% of probability mass lies in $[m/s^3 , ms^3]$
- if you expect η to be around one, choose a value of one for m (and, therefore, $\mu = 0$)
- if you expect η is in the interval [0.1 , 10], choose

$$s = \sqrt[3]{10} \text{ and } \sigma = \log\left(\sqrt[3]{10}\right)$$

Beta prior for ρ

- Bayesian composite criterion

$$D_B = \int_0^1 \log\left(D(\rho)\right) \frac{\Gamma(\kappa+\lambda)}{\Gamma(\kappa)\Gamma(\lambda)} \rho^{\kappa-1}(1-\rho)^{\lambda-1} d\rho$$

- Gauss-Jacobi quadrature

$$D_B \approx \frac{1}{2^{\kappa+\lambda-1}} \cdot \frac{\Gamma(\kappa+\lambda)}{\Gamma(\kappa)\Gamma(\lambda)} \sum_{i=1}^{R} w_i^{GJ} \log\left(D\left(\frac{\alpha_i^{GJ}+1}{2} \right)\right)$$

- R is small

Choice of κ and λ

- expectation of a beta distributed random variable is $\kappa/(\kappa+\lambda)$
- large values for the sum $\kappa+\lambda$ correspond to a smaller variance
- for the expectation $\kappa/(\kappa+\lambda)$ to be 0.5, κ and λ have to be identical
- in case of little prior uncertainty about the value of ρ, the sum $\kappa+\lambda$ has to be large

Example 1

8 runs

4 blocks of size 2

2 factors

full quadratic model

8 parameters

Example 1

Blocks	x_1	x_2
1	-1	1
	1	0
2	0	1
	-1	-1
3	0	-1
	1	1
4	1	-1
	-1	0

Old criterion

Blocks	x_1	x_2
1	1	1
	0	0
2	-1	-1
	-1	-1
3	1	-1
	-1	1
4	1	0
	0	1

New criterion

Example 1
(criterion values)

approach	composite criterion	fixed effects	variance components
any prior mean	-0.39	4.25	-2.2
old	$-\infty$	5.31	$-\infty$

Example 2

8 runs

4 whole plots of size 2

1 whole-plot factor (w)

1 sub-plot factor (s)

full quadratic model

8 parameters

Example 2

WPs	w	s
1	1	-1
	1	1
2	0	-1
	0	0
3	-1	-1
	-1	1
4	-1	1
	-1	0

Old criterion

WPs	w	s
1	1	1
	1	-1
2	-1	-1
	-1	1
3	-1	-1
	-1	1
4	0	-1
	0	0

New criterion

(small prior mean)

WPs	w	s
1	1	1
	1	-1
2	0	-1
	0	0
3	0	1
	0	0
4	-1	1
	-1	-1

New criterion

(large prior mean)

Example 2
(criterion values)

approach	composite criterion	fixed effects	variance components
small prior mean	-0.49	3.64	-2.20
large prior mean	-0.51	3.92	-2.33
old	-0.56	4.04	-2.46

Example 3

24 runs

6 blocks of size 4

5 factors

full quadratic model

23 parameters

Example 3
(average estimates)

η	approach	σ_ε^2	σ_γ^2	est η
10	new	0.985	11.732	10.612
	old	19.496	4.752	0.222
1	new	4.481	6.418	1.389
	old	14.355	2.824	0.220
0.1	new	6.820	1.755	0.301
	old	9.272	0.731	0.217

Example 3
(estimation problems)

η	appr.	zero σ_ε^2	large σ_ε^2	zero σ_γ^2	large σ_γ^2	conv.
10	new	531	21	319	8	1085
	old	753	77	3607	33	1571
1	new	240	12	1764	7	455
	old	874	85	4108	57	848
0.1	new	131	4	3701	1	116
	old	1053	84	4763	53	317

Discussion

- new design criterion with dual objective:
  - fixed-effects estimation
  - variance-component estimation
- Bayesian flavor
  - prior distribution for relative magnitude of variance components
  - quadrature approaches
- fewer estimation problems with variance components
- new approach is not a panacea: designs with few runs and blocks/whole plots will always be problematic

Optimal design of blocked and split-plot experiments for fixed-effects and variance-components estimation

Peter Goos

Kalliopi Mylona

Universiteit Antwerpen