The Robustness of Locally Optimal Experimental Designs in Growth Problems: a Case Study.

M. João Paulo, Dieter Rasch
Wageningen Agricultural University, Sub-department of Mathematics
Dreijenlaan 4
6703 HA Wageningen, The Netherlands
maria.paulo@wts.wk.wau.nl

Margarida Tomé
Instituto Superior de Agronomia, D.E.F.
Tapada da Ajuda
1399 Lisboa Codex, Portugal
magatome@isa.utl.pt

1. Introduction

With an optimal design one wants to determine the ideal allocation of observations for the estimation of the unknown parameter vector \( \theta \) of a given model \( E(y_i) = f(x_i; \theta) \), with \( \theta^T = [\theta_1 \ldots \theta_p]^T \), \( x_i = (x_{i1}, \ldots, x_{in}) \), the \( x_i \) from a given experimental region \( ER \). Let \( V(\theta; X) \) be the asymptotical variance-covariance matrix of the least squares estimator \( \hat{\theta} \) of \( \theta \). We consider the D-optimality criterion based on the determinant of \( V(\theta; X) \) so that the D-optimal design is defined as that \( X^* \) for which

\[
(1) \quad k(\theta, X^*) = \min_x \left\{ \det \left( \frac{1}{\sigma^2}V(\theta; X) \right) \right\}.
\]

If \( f \) is intrinsically non-linear (Rasch, 1990), \( V(\theta; X) \) depends not only on \( X \) but also on \( \theta \) and thus the D-optimal design will also depend on \( \theta \). Such designs are therefore called locally D-optimal.

By the material of a case study we investigate the robustness of the D-optimal design against miss-specification of the parameter \( \theta \). For this we used growth data of the diameter of 24 cork oaks from Portuguese forests. First we had to find a function that could fit the diameter growth curves of each of the 24 trees. Then we used the D-optimal design for each of the trees and looked at its robustness (analogue to Rasch, 1995) against the choice of the parameters of each of the other trees. This work has been financially funded by program PRAXIS XXI.

2. Materials and methods

Measurements of the annual diameter growth of each of 24 cork oaks with ages between 41 and 139 years (Tomé, 1998) were used. We fitted 10 growth functions on the measurements from each tree. Amongst these functions were the exponential, the logistic, the Gompertz and the Bertalanffy. We used the residual variance criterion to rank the ten functions for each tree. First we had to find a function that could fit the diameter growth curves of each of the 24 trees. Then we used the D-optimal design for each of the trees and looked at its robustness (analogue to Rasch, 1995) against the choice of the parameters of each of the other trees. This work has been financially funded by program PRAXIS XXI.
(2) \[ R = \frac{k(\hat{\theta}_i, X_i^\prime)}{k(\hat{\theta}_i, X_s^\prime)}, \]

3. Results

The growth curves of the trees were very different. The Bertalanffy function provided a good fit to every tree, having always the first or second lowest residual variance. The other functions fitted poorly. We decided therefore to use this function, given by

\[ y_{ij} = \left( \alpha_i + \beta_i e^{r_{ij}} \right) \]

to model the diameter growth. Here \( y_{ij} \) is the measurement in point \( j \) for individual \( i \) and \( \theta^T = (\alpha, \beta, \gamma) \). We obtained very different D-optimal designs for the 24 growth curves. However if we use the optimal design of another set of parameters we still have a higher robustness than with any 12-point design in 60% of the cases. We found that the 3-point design with the support points (5, 25, 150) was better than any 12-point design and had \( R>0.1 \) for all except one of the trees.

4. Discussion

Optimal designs provide an economic and efficient way to estimate unknown parameters of a growth curve. We wanted to find the optimal design for a problem in forestry. However the individual behaviour of trees is quite variable, and thus we could not find one design that fitted every tree in the study equally well. We have one optimal design for each tree, which is not useful in practice. We found however that optimal designs under parameter mis-specification are still better than 12-point designs in most cases. Further, the 3-point design (5, 25, 150) seems to perform almost as good as the optimal designs and clearly better than the 12-point designs tested. The result agrees with previously published work (Rasch et al, 1997). A question however remains: do we want to wait 150 years to estimate a growth curve or can we use a shorter \( ER \) to estimate the parameters with good precision?

REFERENCES


Rasch, D Hendrix, E. and Boer, E.; (1997). Replicationfree optimal designs in regression analysis, Comp. Statist. 12, 19-52


RÉSUMÉ

Nous avons pris des données du diamètre de 24 arbres et avons recherché les possibilités d'applications des dessins D-optimales à la croissance du diamètre des arbres. Nous avons étudié la robustesse contre le choix des paramètres de chacun des autres arbres. Nous avons trouvé un bon dessin pour tous les arbres, mais il y a encore des questions qui se lèvent de cette investigation.