Design of Nonlinear Univariate Calibration with Application to Immunoassay

Johannes Forkman

Research Report
Centre of Biostochastics

Design of Nonlinear Univariate Calibration with Application to Immunoassay

JOHANNES FORKMAN[1]

Centre of Biostochastics
Swedish University of Agricultural Sciences
P.O. Box 7032, SE-750 07 Uppsala, Sweden

Abstract

A method is proposed for designing nonlinear univariate calibration of measuring instruments. The problem addressed is how to select a set of design points (standards or calibrators) in order to minimize the errors in the inverse predictions. The method utilizes knowledge not only about the expected values of the curve parameters, but also about the covariance matrix. A design criterion is suggested for analytical procedures, according to which the coefficient of variation and the area under the precision profile are minimized.

Keywords: Optimal design, Nonlinear regression, Inverse prediction, Dose-Response, Analytical procedure, Immunoassay

---

[1]E-mail address to the correspondence author: johannes.forkman@bt.slu.se

1 Introduction

An instrument measuring a physical quantity is calibrated when the relationship between its response and the physical quantity is established. Mathematically the calibration curve, established by the calibration, is a function from a physical quantity to a response. The physical quantity is the independent variable, and the response is the dependent variable. The calibration curve is fitted to measurements of entities with known physical quantities. These entities are often called standards or calibrators. The known physical quantities are called design points.

Calibration curves are for example used in many analytical procedures for determining the concentration of an analyte in a biological sample. The calibrators are samples with known concentrations, and the design points are the concentrations of the calibrators.

Calibrated measuring instruments are used for inverse prediction. Responses are translated into estimates of physical quantities via the calibration curve. The problem discussed in this article is how to select the design points in such a way that the inverse predictions, i.e. the estimates of the independent variable, are optimized with respect to a design criterion.

The errors in the inverse predictions are dependent on the parameters of the calibration curve. This makes the search for an optimal design difficult, because the parameters are not known in advance. Indeed, the calibration is performed in order to estimate the unknown current parameter values.

Since the parameters are not known designs are often optimized locally, i.e. conditioned on some a priori decided fixed parameter values that are considered typical for the method. A second approach is to assume that the parameters belong to a pre-specified multivariate distribution and look for a Bayesian optimal design. Such were reviewed by Chaloner and Verdinelli (1995). As explained by François et al. (2004) the Bayesian approach "is not conceptually difficult, but computer intensive because it requires a multidimensional integration over the parameters space at each cycle of the design search." They argued that the first approach, to search for a locally optimal design, "is quite realistic since calibration is an iterative process where the researcher often has some knowledge on the expected model parameters." Kuljus et al. (2006), on the other hand, argued in a related problem that since the parameters are unknown there is not much use of a local optimization.

In this article a third approach is suggested, which is more complex than the first approach, but perhaps more convenient than the second. Knowledge often exists not only about the expected values of the parameters, but also about the variances. A measuring instrument manufacturer may know typical

parameters values, but also that the parameter values vary, dependent on e.g. temperature or humidity, and that some parameters are correlated. This variation could be described by a covariance matrix $\mathbf{\Sigma}$. The method proposed in this article utilizes knowledge not only of the model parameters' expected values, but also of their covariance matrix $\mathbf{\Sigma}$. It is, however, not necessary to specify an exact multivariate distribution for the parameters. The idea is to consider the curve parameters as random, and to estimate the unconditional variance and bias in the inverse prediction by linear approximations.

The covariance matrix of the parameters should not be confused with the covariance matrix of the parameter estimates, which is dependent on a specific parameter vector $\boldsymbol{\beta}$. When the function is linear in the parameters the covariance matrix of the parameter estimates is $\sigma^2(\mathbf{X'X})^{-1}$, where $\mathbf{X}$ is the design matrix, provided that the errors are uncorrelated with constant variance σ^2. A reasonable strategy in optimization problems is to search for a design that minimizes the determinant of $(\mathbf{X'X})^{-1}$ (Atkinson and Donev 1992). When the function is nonlinear in the parameters, which is often the case in analytical procedures, the covariance matrix can be approximated by $\sigma^2(\mathbf{F'F})^{-1}$, where $\mathbf{F}$ denotes the matrix of partial derivatives of the function with respect to the parameters. The determinant of $(\mathbf{F'F})^{-1}$ is a common design criterion in nonlinear regression, suggested by Box and Lucas (1959). We shall instead minimize the errors in the inverse predictions, as suggested for linear calibration by Ott and Myers (1968) and Buonaccorsi (1986). Ott and Myers (1968) considered minimization of the integral over the measuring range of the mean square error. Buonaccorsi (1986) focused on the asymptotic variance. Rocke and Jones (1997) considered the problem of selecting design points for immunoassays using the four-parameter logistic function. They suggested maximization of the reciprocal of the asymptotic variance over the measuring range on the logarithmic scale and suggested a method for the case that the calibrators are manufactured by serial dilution with constant dilution ratios. François et al. (2004) studied the design problem of nonlinear calibration in a more general setting. They recommended minimization of the integral over the measuring range of the asymptotic variance in the inverse predictions. We shall, for analytical procedures, propose that the precision profile is used as design criterion, based on the coefficient of variation.

2   Method

The model is described in Section 2.1. Calculations relevant for locally optimal designs are presented in Section 2.2. The new method, based on random

curve parameters, is described in Section 2.3. The design criterion for the optimization is considered in Section 2.4.

2.1 The model

Assume that the calibration curve $f = f(\xi, \boldsymbol{\beta})$, with parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)$, is a strictly monotonic continuous function of the physical quantity ξ. Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$ be a vector of m design points. Let $f_i(\boldsymbol{\beta}) = f(\xi_i, \boldsymbol{\beta})$ and $\mathbf{f}(\boldsymbol{\beta}) = (f_1(\boldsymbol{\beta}), f_2(\boldsymbol{\beta}), \ldots, f_m(\boldsymbol{\beta}))'$. Let $\mathbf{F}(\boldsymbol{\beta})$ denote the $m \times p$ matrix of partial derivatives $\partial f_i(\boldsymbol{\beta})/\partial \beta_j$. Let $\mathbf{y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\epsilon}$, where the error vector $\boldsymbol{\epsilon}$ has expected value $\mathbf{0}$ and diagonal covariance matrix $\mathbf{D}$. Assume that the rank of $\mathbf{F}(\boldsymbol{\beta})$ is p and, throughout this article, that all required derivatives and moments exist.

Let ξ, without index, denote the unknown parameter that shall be estimated by inverse prediction (e.g. the concentration of a sample). We observe a response $y = \mu + e$, where $\mu = f(\xi, \boldsymbol{\beta})$, and assume that e varies, independently of $\boldsymbol{\epsilon}$, with expected value 0 and variance σ^2.

The generalized least squares estimate of $\boldsymbol{\beta}$ is the $\mathbf{b}$ that minimizes the generalized least squares, i.e.

$$(\mathbf{y} - \mathbf{f}(\mathbf{b}))'\mathbf{D}^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{b})) = \min_{\boldsymbol{\beta}} \ (\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}))'\mathbf{D}^{-1}(\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})). \tag{1}$$

It is well known that $\mathbf{b}$ is the ordinary least squares estimate in a transformed model $\mathbf{C}^{-1}\mathbf{y} = \mathbf{C}^{-1}\mathbf{f}(\boldsymbol{\beta}) + \mathbf{C}^{-1}\boldsymbol{\epsilon}$, where $\mathbf{C}$ is such that $\mathbf{D} = \mathbf{C}\mathbf{C}'$. Jennrich (1969) showed that $\mathbf{b}$ exists, under mild regularity conditions, and studied asymptotic properties. If the number of distinct design points is finite, then $\mathbf{b}$ is consistent as the number of replicates per design point is increasing (Malinvaud 1970; Gallant 1975). Given some further regularity conditions (Seber and Wild 1989), $\mathbf{b}$ is asymptotically normally distributed with expected value $\boldsymbol{\beta}$ and covariance matrix

$$\mathbf{V}(\boldsymbol{\beta}) = (\mathbf{F}(\boldsymbol{\beta})'\mathbf{D}^{-1}\mathbf{F}(\boldsymbol{\beta}))^{-1}, \tag{2}$$

where we write $\mathbf{V}(\boldsymbol{\beta})$ to indicate that the covariance matrix $\mathbf{V}$ of $\mathbf{b}$ is dependent on $\boldsymbol{\beta}$.

Provided that $\mathbf{b}$ is close to $\boldsymbol{\beta}$

$$\mathbf{f}(\mathbf{b}) \approx \mathbf{f}(\boldsymbol{\beta}) + \mathbf{F}(\boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta}). \tag{3}$$

Substituting (3) in (1), yields $(\boldsymbol{\epsilon} - \mathbf{F}(\boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta}))'\mathbf{D}^{-1}(\boldsymbol{\epsilon} - \mathbf{F}(\boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta}))$, which is minimized when $\mathbf{b} = \boldsymbol{\beta} + \mathbf{F}^{-1}(\boldsymbol{\beta})\boldsymbol{\epsilon}$. Thus, when $\mathbf{b}$ varies near $\boldsymbol{\beta}$, the variance in $\mathbf{b}$ approximately equals (2) even if the number of design points is small.

The physical quantity ξ is estimated by $x = f^{-1}(y, \mathbf{b})$, where f^{-1} denotes the inverse of f considered as a function of ξ. In a neighborhood of $(\mu, \boldsymbol{\beta})$,

$$x = \xi + k(\boldsymbol{\beta})(y - \mu) + \mathbf{g}'(\boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta}) + O(\delta^2), \qquad (4)$$

where δ is the radius of the neighborhood and

$$k(\boldsymbol{\beta}) = \left. \frac{\partial f^{-1}}{\partial y} \right|_{(y,\mathbf{b})=(\mu,\boldsymbol{\beta})}, \qquad \mathbf{g}(\boldsymbol{\beta}) = \left. \frac{\partial f^{-1}}{\partial \mathbf{b}} \right|_{(y,\mathbf{b})=(\mu,\boldsymbol{\beta})}. \qquad (5)$$

Notice, as will be important in Section 2.3, that $k(\boldsymbol{\beta})$ and $\mathbf{g}(\boldsymbol{\beta})$, as defined by (5), can be calculated for a given ξ by

$$k(\boldsymbol{\beta}) = \left. \left(\frac{\partial f}{\partial x} \right)^{-1} \right|_{(x,\mathbf{b})=(\xi,\boldsymbol{\beta})}, \qquad (6)$$

$$\mathbf{g}(\boldsymbol{\beta}) = \left. -\frac{\partial f}{\partial \mathbf{b}} \left(\frac{\partial f}{\partial x} \right)^{-1} \right|_{(x,\mathbf{b})=(\xi,\boldsymbol{\beta})}. \qquad (7)$$

2.2 Fixed curve parameters

In this subsection we study the variance and the expected value of the inverse prediction x of ξ, conditioned on $\boldsymbol{\beta}$.

Under assumption that the variance of the last term in (4) is small

$$Var(x|\boldsymbol{\beta}) \approx k^2(\boldsymbol{\beta})\sigma^2 + \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta}), \qquad (8)$$

since y and $\mathbf{b}$ are independent. The first term in (8) is the variance σ^2 of the observation in response transformed via linear approximation into variance in inverse prediction. The second term is the asymptotic covariance matrix $\mathbf{V}(\boldsymbol{\beta})$ of the curve parameter estimates transformed via linear approximation into variance in inverse prediction. The approximate conditional variance in x is thus composed of two parts, one caused by the variation in the measurement of the entity with unknown ξ, and one caused by the variation in the p measurements of the design points.

The bias in nonlinear regression is known to be an order of magnitude smaller than the standard errors of the parameter estimates (Box 1971). Let $\mathbf{A}_i(\boldsymbol{\beta})$ denote the $p \times p$ matrix with elements $\partial^2 f_i(\boldsymbol{\beta})/(\partial \beta_r \partial \beta_s)$, $i = 1, 2, \ldots, m$. Box (1971) showed that, with the additional assumption of normally distributed measurements, $E[\mathbf{b} - \boldsymbol{\beta}] \approx \mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})$, where the elements

of the vector $\mathbf{z}(\boldsymbol{\beta})$ are $z_i(\boldsymbol{\beta}) = -\text{tr}\{\mathbf{V}(\boldsymbol{\beta})\mathbf{A}_i(\boldsymbol{\beta})\}/2$, $i = 1, 2, \ldots, m$. Thus, by (4),

$$E(x|\boldsymbol{\beta}) \approx \xi + \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta}), \tag{9}$$

and approximately the bias equals $c(\boldsymbol{\beta}) = \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})$.

2.3   Random curve parameters

In Section 2.2 we assumed that $\boldsymbol{\beta}$ was a fixed known parameter vector. We now extend the theory by assuming that $\boldsymbol{\beta}$ is a random parameter vector with expected value $\boldsymbol{\beta}_0$ and covariance matrix $\boldsymbol{\Sigma}$. We study the variance and the expected value of the inverse prediction given known $\boldsymbol{\beta}_0$ and $\boldsymbol{\Sigma}$.

The unconditional variance includes two terms, the mean of the conditional variance and the variance in the conditional mean. The variance in the inverse prediction is thus, by (8) and (9),

$$\begin{aligned}
Var(x) &= E(Var(x|\boldsymbol{\beta})) + Var(E(x|\boldsymbol{\beta})) \\
&\approx \sigma^2 E(k^2(\boldsymbol{\beta})) + E(\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta})) \\
&\quad + Var(\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}^{-1}(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})), \tag{10}
\end{aligned}$$

where calculation of $k(\boldsymbol{\beta})$ and $\mathbf{g}(\boldsymbol{\beta})$ for a given ξ is made possible by (6) and (7) respectively. The third term on the right hand side in (10) is the variance in the bias, which is often very small. For this reason we focus on $\sigma^2 E(k^2(\boldsymbol{\beta}))$ and $E(\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta}))$ and make Taylor series expansions of $k^2(\boldsymbol{\beta})$ and $\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta})$, as functions of $\boldsymbol{\beta}$, about $\boldsymbol{\beta} = \boldsymbol{\beta}_0$. To make notation easier we write $\mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta})$ instead of $\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta})$. Referring to Kollo and von Rosen (2005, p. 152),

$$\begin{aligned}
k^2(\boldsymbol{\beta}) &= k^2(\boldsymbol{\beta}_0) + (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \left.\frac{dk^2}{d\boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
&\quad + \frac{1}{2}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 2}\text{vec}\left(\frac{d^2 k^2}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} + r, \tag{11}
\end{aligned}$$

where in a neighborhood of $\boldsymbol{\beta}_0$, for some $\boldsymbol{\rho}$ in the neighborhood, the error term r is

$$r = \frac{1}{6}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 3}\text{vec}\left(\frac{d^3 k^2}{d\boldsymbol{\beta}^3}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\rho}}.$$

The vectorization operator vec stacks the columns of the the matrix, one under the other. The Kroneckerian power $\mathbf{A}^{\otimes k}$ of $\mathbf{A}$ is defined in Appendix A.

Similarly,

$$
\begin{aligned}
\mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}) \;=\; & \mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}_0) + (\boldsymbol{\beta} - \boldsymbol{\beta}_0)'\, \frac{d(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
& + \frac{1}{2}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 2}\mathrm{vec}\left(\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} + s, \qquad (12)
\end{aligned}
$$

where in a neighborhood of $\boldsymbol{\beta}_0$, for some $\boldsymbol{\varsigma}$ in the neighborhood, the error term s is

$$
s = \frac{1}{6}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 3}\mathrm{vec}\left(\frac{d^3(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^3}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\varsigma}}.
$$

By (10), (11) and (12),

$$
\begin{aligned}
Var(x) \;\approx\; & \sigma^2 k^2(\boldsymbol{\beta}_0) + \mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}_0) + \frac{\sigma^2}{2}\mathrm{vec}'\boldsymbol{\Sigma}\,\mathrm{vec}\left(\frac{d^2 k^2}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
& + \frac{1}{2}\,\mathrm{vec}'\boldsymbol{\Sigma}\,\mathrm{vec}\left(\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}, \qquad (13)
\end{aligned}
$$

provided that the variance in the bias, i.e. the last term in (10), could be neglected. The $p \times p$ matrix $d^2(\mathbf{g}'\mathbf{V}\mathbf{g})/d\boldsymbol{\beta}^2$ in (13) can be calculated via the first and second order derivatives of $\mathbf{g}(\boldsymbol{\beta})$ and $\mathbf{F}(\boldsymbol{\beta})$. The formulas are given in Appendix B.

The unconditional expectation is, by (9),

$$
E(x) = E(E(x|\boldsymbol{\beta})) \approx E(\xi + \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})).
$$

By a Taylor series expansion of the bias $c(\boldsymbol{\beta}) = \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})$ about $\boldsymbol{\beta} = \boldsymbol{\beta}_0$

$$
\begin{aligned}
c(\boldsymbol{\beta}) \;=\; & c(\boldsymbol{\beta}_0) + (\boldsymbol{\beta} - \boldsymbol{\beta}_0)'\, \frac{dc}{d\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
& + \frac{1}{2}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 2}\mathrm{vec}\left(\frac{d^2 c}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} + t, \qquad (14)
\end{aligned}
$$

where in a neighborhood of $\boldsymbol{\beta}_0$, for some $\boldsymbol{\tau}$, the error term t is

$$
t = \frac{1}{6}((\boldsymbol{\beta} - \boldsymbol{\beta}_0)')^{\otimes 3}\mathrm{vec}\left(\frac{d^3 c}{d\boldsymbol{\beta}^3}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\tau}},
$$

(Kollo and von Rosen, 2005). The expectation of the second term in (14) is zero and the following terms are often small. Then

$$
E(x) \approx \xi + \mathbf{g}'(\boldsymbol{\beta}_0)\mathbf{V}(\boldsymbol{\beta}_0)\mathbf{F}'(\boldsymbol{\beta}_0)\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta}_0). \qquad (15)
$$

6

2.4 Optimization

We have in Section 2.2 and 2.3 seen how to calculate the variance and the expected value of the inverse prediction of the unknown quantity ξ. We now want to find an optimal set of design points that minimizes the variance and the bias in the inverse prediction. Generally we want to minimize a function h_ξ of the variance and the expected value. This function could be the variance or the mean square error, but it could also be the coefficient of variation γ as defined by

$$\gamma = \sqrt{Var(x)}/E(x). \tag{16}$$

The coefficient of variation (16) is often used in analytical procedures and is recommended e.g. by De Silva et al. (2003).

If $\boldsymbol{\beta}$ is considered as fixed and the search, for a locally optimal design, is made given a known $\boldsymbol{\beta}$ equation (8) and (9) could be used for calculating h_ξ. In this article we propose that $\boldsymbol{\beta}$ is instead modeled as a random variable, and that h_ξ is calculated by

$$
\begin{aligned}
Var(x) \approx{} & \sigma^2 k^2(\boldsymbol{\beta}_0) + \mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}_0) + \frac{\sigma^2}{2}\text{vec}'\boldsymbol{\Sigma}\ \text{vec}\left(\frac{d^2 k^2}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
& + \frac{1}{2}\ \text{vec}'\boldsymbol{\Sigma}\ \text{vec}\left(\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2}\right)'\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0},
\end{aligned}
$$

from (13), and

$$E(x) \approx \xi + \mathbf{g}'(\boldsymbol{\beta}_0)\mathbf{V}(\boldsymbol{\beta}_0)\mathbf{F}'(\boldsymbol{\beta}_0)\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta}_0),$$

from (15).

In practice ξ is unknown, because it is the unknown quantity that shall be measured by the instrument. For this reason we want to minimize the function h_ξ for all ξ that belong to the measuring range. The design criterion, which we want to minimize over $\boldsymbol{\xi}$, is consequently

$$\int h_\xi \pi_\xi\ d\xi, \tag{17}$$

where π is a prior probability density function of ξ. Typically we assume that ξ belongs to a measuring range $[\alpha, \omega]$, where often α is the lowest design point and ω is the largest. We then want to minimize the function h_ξ for all ξ, with weights defined by the prior distribution π_ξ, which takes the value 0 outside $[\alpha, \omega]$.

We shall in the following example let h_ξ equal the coefficient of variation, estimated by $\sqrt{Var(x)}/E(x)$ as calculated by (13) and (15) under assumption of an expected parameter vector $\boldsymbol{\beta}_0$ and a covariance matrix $\boldsymbol{\Sigma}$.

7

3 Example

ImmunoCAP ECP (Phadia AB, Uppsala, Sweden) is an immunoassay method for monitoring inflammation in asthma. It measures concentration of eosinophil cationic protein in serum. The variance in the measuring errors needs to be small when repeated measurements, perhaps including several calibrations, are made on the same subject. The measuring range of Immuno-CAP ECP is $2 - 200\ \mu$g/l. Calibration is made with the four-parameter logistic function

$$f(\xi, \boldsymbol{\beta}) = \beta_2 + \frac{\beta_1 - \beta_2}{1 + (\xi/\beta_3)^{\beta_4}}, \tag{18}$$

which describes a sigmoid curve when ξ is set off on logarithmic scale. This function is often used for calibration of analytical procedures. The parameter β_1 is the expected response at concentration $\xi = 0\ \mu$g/l, and β_2 is the limit of f when $\xi \to \infty$. When ξ equals β_3 the expected response f equals $(\beta_1 + \beta_2)/2$. The slope of the curve is controlled by β_4. The inverse of the four-parameter logistic function is, by (18),

$$f^{-1}(y, \boldsymbol{\beta}) = \beta_3 \left(\frac{y - \beta_1}{\beta_2 - y} \right)^{\frac{1}{\beta_4}}.$$

We let $\widehat{\gamma}$, as an estimate of the coefficient of variation (16), be defined as the square root of the right hand side of (13), divided by the right hand side of (15). We use as design criterion

$$\int_2^{200} h_\xi \pi_\xi\ d\xi = \frac{1}{\log(100)} \int_{\log 2}^{\log 200} \widehat{\gamma}\ d(\log \xi), \tag{19}$$

where $h_\xi = \widehat{\gamma}$ and $\pi_\xi = (\xi \log 100)^{-1}$. With this criterion the coefficient of variation is averaged over the measuring range on logarithmic scale. In a precision profile the coefficient of variation is plotted against concentration ξ over the measuring range (Dudley et al. 1985). It is sensible to display the measuring range on the logarithmic scale since subjects are often more evenly distributed on the logarithmic scale than on the original scale and since the logistic function (18) represents a logistic regression of the proportion $(\mu - \beta_1)/(\beta_2 - \beta_1)$ on $\log \xi$. By use of the design criterion (19) the area under the precision profile on the logarithmic scale is minimized. For a given set of design points, we calculate the integral (19) numerically by the trapezoidal rule.

The calibration curve (18) is fitted by generalized least squares. In our search for an optimal design we assume that the diagonal covariance matrix

8

D has diagonal elements $\phi(f_i(\boldsymbol{\beta}_0))^\theta$, i = 1,2,...,4, and that $\sigma^2 = \phi\mu^\theta$, where $\phi = 0.00067$, $\theta = 1.88$ and $\boldsymbol{\beta}_0 = (40, 34000, 150, 1.4)'$. We assume that $\boldsymbol{\beta}$ is a random vector with expected value $\boldsymbol{\beta}_0$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 100 & -7680 & -80 & 2.4 \\ -7680 & 10240000 & 12800 & -900 \\ -80 & 12800 & 400 & -0.64 \\ 2.4 & -900 & 0.64 & 0.16 \end{pmatrix}.$$

To prevent extrapolation outside the measuring range we fix the lowest design point to 2 μg/l and the highest to 200 μg/l. We search for an optimal set of 3 intermediate design points, using MATLAB 6.5 (The Mathworks Inc., Natick, MA, USA) and the function fminsearch. This function makes use of the simplex search method by Nelder and Mead (1965) and algorithms given by Lagarias, Reeds, Wright and Wright (1998). In each step we let MATLAB perform the quadrature based on 1 000 points equally spaced on the logarithmic scale. With starting value $(2.00, 10.0, 50.0, 100, 200)'$ the search algorithm converges after 93 iterations at $\mathbf{x} = (2.00, 6.66, 18.0, 82.3, 200)'$, where the integral (19) equals 2.8389%. Rounded into $(2.00, 7.00, 18.0, 80.0, 200)'$ the integral is only slightly larger: 2.8392%.

Figure 1 is the precision profile for this choice of design points. The solid line in Figure 1 is the calculated coefficient of variation $\widehat{\gamma}$ under assumption of random parameters. The bias, as calculated by the last term of (15) and expressed as a percentage of concentration, is for comparison included in Figure 1 as a dotted line.

If we instead consider $\boldsymbol{\beta} = (40, 34000, 150, 1.4)'$ as a fixed parameter vector, and calculate $\widehat{\gamma}$ as the square root of the right hand side of (8), divided by the right hand side of (9), the algorithm converges after 98 iterations at the locally optimal design $\mathbf{x} = (2.00, 5.70, 13.2, 60.2, 200)'$, where the integral (19) equals 1.9727%.

4 Discussion

A locally optimal design is a design that is optimized for a given fixed parameter vector. There is no guarantee that a locally optimal design gives small errors also for other parameter vectors. This is a problem since the parameters are likely to change. Otherwise it would not be necessary to calibrate the instrument. We must require of a good design for calibration that it performs well under varying curve parameters.
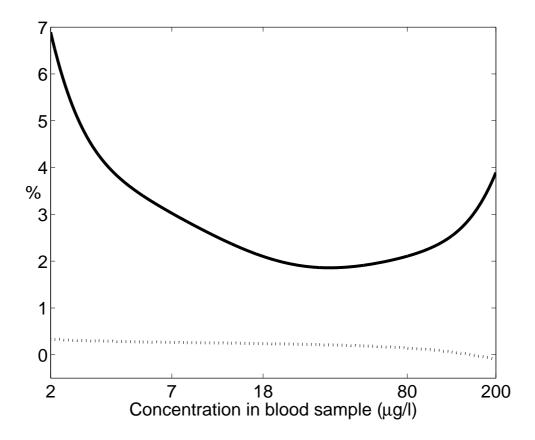
Figure 1: Precision profile for the set of design points $\mathbf{x} = (2.00, 7.00, 18.0, 80.0, 200)'$. The integral of the coefficient of variation (solid line) is minimized over the measuring range on logarithmic scale. The bias (dotted line), illustrated as a percentage of concentration, is small in comparison with the standard deviation.

Section 2.2 included approximations useful for the search of a locally optimal design, similar as those presented by Rocke and Jones (1997) and François et al. (2004). The equations of Section 2.2 were extended in Section 2.3 to make possible a search basad on the assumption of a random parameter vector.

In order to model the more correct assumption of random parameters not only the expected parameter values are needed but also the variances and covariances. These could be estimated from a dataset of measurements performed under varying conditions. An instrument manufacturer could for example perform precision studies in countries with different climatic conditions and estimate the variation in the calibration curves. It is more difficult to make the assumptions required by bayesian optimization about the exact mathematical form of the multivariate distribution of the parameters. The method introduced in this article, based on the equations given in Section 2.3, only requires the expected values and the covariance matrix.

The minimization of the area under the precision profile is closely related to the average variance criterion, called V-optimality by Atkinson and Donev (1992), AV-optimality by Buonaccorsi (1986) and, to emphasize the application on inverse prediction, V_I-optimality by François et al. (2004).

Many authors have studied the problem of designing non-linear regression under the assumption of normally distributed measurements. It should be noted, however, that the assumption of normality, in this article, is only needed for the estimation of the bias. As illustrated by the example the bias in the inverse prediction is often small compared with the standard deviation. It could thus be acceptable to simplify the method by neglecting the bias and build the design criterion on a function of the variance alone.

Usually it is also assumed, e.g. by François et al. (2004), that the inverse calibration function has an analytical form. According to Schwenke and Milliken (1991) a confidence interval based on $\mathbf{V}(\boldsymbol{\beta})$, as defined by (2), requires the existence of a closed form for the inverse of the nonlinear function. However, equation (7) can be used for differentiation of the inverse function when no closed form exists.

When the function include many parameters a large number of first and second order derivatives need to be calculated. A software package for symbolic mathematics makes it easier. Sometimes search algorithms, such as the simplex method used in the example of Section 3, do not find the global minimum, but stop at a local minimum. For this reason it could be wise to begin with a grid search over the parameter space.

References

[1] Atkinson, A.C. and Donev, A.N. (1992), *Optimum Experimental Designs*, New York, Oxford University Press.

[2] Box, G.E.P. and Lucas, H.L. (1959). Design of experiments in non-linear situations, *Biometrika*, **46**, 77-90.

[3] Box, M.J. (1971). Bias in nonlinear estimation. *Journal of the Royal Statistical Society: Series B*, **32**, 171-190.

[4] Buonaccorsi, J.P. (1986). Design considerations for calibration. *Technometrics*, **28**, 149-155.

[5] Chaloner, K. and Verdinelli, I. (1995). Bayesian Experimental Design: A Review. *Statistical Science*, **10**, 237-304.

[6] DeSilva, B., Smith, W., Weiner, R., Kelley, M., Smolec, J., Lee, B., Khan, M., Tacey, R., Hill, H., and Celniker A. (2003). Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharmaceutical Research*, **20**, 1885-1900.

[7] Dudly, R.A., Edwards, P., Ekins, R.P., Finney, D.J., McKenzie, I.G.M., Raab, G.M., Rodbard, D., and Rodgers, R.P.C. (1985). Guidelines for immunoassay data processing. *Clinical Chemistry*, **31**, 1264-1271.

[8] François, N., Govaerts, B., and Boulanger, B. (2004). Optimal Designs for Inverse Prediction in Univariate Nonlinear Calibration Models. *Chemometrics and Intelligent Laboratory Systems*, **74**, 283-292.

[9] Gallant, A.R. (1975). The power of the likelihood ratio test of location in nonlinear regression models. *Journal of the American Statistical Association*, **70**, 198-203.

[10] Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics*, **40**, 633-643.

[11] Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*, Dordrecht, Springer.

[12] Kuljus, K., von Rosen, D., Sand, S., and Victorin, K. (2006). Comparing experimental designs for benchmark dose calculations for continuous endpoints. *Risk Analysis*, **26**, 1031-1043.

[13] Lagarias, J.C., Reeds, J.A., Wright, M.H., and Wright, P.E. (1998). Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, **9**, 112-147.

[14] Malinvaud, E. (1970). The consistency of nonlinear regressions. *The Annals of Mathematical Statistics*, **41**, 956-969.

[15] Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, **7**, 308-313.

[16] Ott, R.L. and Myers, R.H. (1968). Optimal experimental designs for estimating the independent variable in regression. *Technometrics*, **10**, 811-823.

[17] Rocke, D.M. and Jones, G. (1997). Optimal design for elisa and other forms of immunoassay. *Technometrics*, **39**, 162-170.

[18] Seber, G.A.F. and Wild, C.J. (1989). *Nonlinear Regression*, New York, Wiley.

Appendix A: Direct product

The Kronecker product $\mathbf{A} \otimes \mathbf{B}$ of the $p \times q$ matrix $\mathbf{A}$ and the $m \times n$ matrix $\mathbf{B}$ is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \ldots & a_{1q}\mathbf{B} \\ \vdots & & \vdots \\ a_{p1}\mathbf{B} & \ldots & a_{pq}\mathbf{B} \end{pmatrix}.$$

The Kroneckerian power is given by

$$\mathbf{A}^{\otimes k} = \underbrace{\mathbf{A} \otimes \ldots \otimes \mathbf{A}}_{k \text{ times}}.$$

Appendix B: Matrix derivatives

Following Kollo and von Rosen (2005), define the matrix derivative of order k of $\mathbf{Y}$ by $\mathbf{X}$ as

$$\frac{d^k \mathbf{Y}}{d\mathbf{X}^k} = \frac{d}{d\mathbf{X}}\left(\frac{d^{k-1}\mathbf{Y}}{d\mathbf{X}^{k-1}}\right),$$

where

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\text{vec}'\mathbf{Y}}{d\text{vec}\mathbf{X}}.$$

According to this definition

$$\mathbf{F}(\boldsymbol{\beta}) = \left(\frac{d\mathbf{f}}{d\boldsymbol{\beta}}\right)'.$$

Let $\mathbf{I}_p$ denote the $p \times p$ identity matrix, and let the partitioned matrix $\mathbf{K}_{p,q}$ denote the $pq \times pq$ commutation matrix, in which the (j, i)–th element in block (i, j) equals one, and all the other elements in that block equal zero.

By application of the rules for matrix differentiation (Kollo and von Rosen 2005, p. 149)

$$\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2} = 2\left(\frac{d^2\mathbf{g}}{d\boldsymbol{\beta}^2}\,(\mathbf{V}\mathbf{g}\ \otimes\ \mathbf{I}_p)\ +\ \frac{d(\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}}\left(\frac{d\mathbf{g}}{d\boldsymbol{\beta}}\right)'\right)$$
$$-\ \frac{d^2(\mathbf{F}'\mathbf{D}^{-1}\mathbf{F})}{d\boldsymbol{\beta}^2}\left((\mathbf{V}\mathbf{g})^{\otimes 2}\ \otimes\ \mathbf{I}_p\right)$$
$$-\left(\left(\frac{d(\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}}\ \otimes\ \mathrm{vec}'(\mathbf{V}\mathbf{g})\right)\right.$$
$$\left.+\left(\mathrm{vec}'(\mathbf{V}\mathbf{g})\ \otimes\ \frac{d(\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}}\right)\right)\left(\frac{d(\mathbf{F}'\mathbf{D}^{-1}\mathbf{F})}{d\boldsymbol{\beta}}\right)',$$

where

$$\frac{d(\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}} = \frac{d\mathbf{g}}{d\boldsymbol{\beta}}\mathbf{V}\ -\ \frac{d(\mathbf{F}'\mathbf{D}^{-1}\mathbf{F})}{d\boldsymbol{\beta}}\,\mathbf{V}^{\otimes 2}\,(\mathbf{g}\ \otimes\ \mathbf{I}_p),$$

$$\frac{d(\mathbf{F}'\mathbf{D}^{-1}\mathbf{F})}{d\boldsymbol{\beta}} = \frac{d\mathbf{F}}{d\boldsymbol{\beta}}\,\mathbf{K}_{p,m}((\mathbf{D}^{-1}\mathbf{F})\otimes\ \mathbf{I}_p) + \frac{d\mathbf{F}}{d\boldsymbol{\beta}}\,(\mathbf{I}_p\otimes(\mathbf{D}^{-1}\mathbf{F}))$$

and

$$\frac{d^2(\mathbf{F}'\mathbf{D}^{-1}\mathbf{F})}{d\boldsymbol{\beta}^2} = \frac{d^2\mathbf{F}}{d\boldsymbol{\beta}^2}\,(\mathbf{K}_{p,m}\otimes\ \mathbf{I}_p)((\mathbf{D}^{-1}\mathbf{F})\otimes\ \mathbf{I}_{p^2})$$
$$+\left(\left(\frac{d\mathbf{F}}{d\boldsymbol{\beta}}\,(\mathbf{I}_p\otimes\ \mathbf{D}^{-1})\right)\otimes\ \mathrm{vec}'\mathbf{I}_p\right)$$
$$\left(\mathbf{I}_p\otimes\mathbf{K}_{m,p}\otimes\ \mathbf{I}_p\right)\left(\mathbf{I}_{p^2}\otimes\left(\frac{d\mathbf{F}}{d\boldsymbol{\beta}}\,\mathbf{K}_{p,m}\right)'\right)$$
$$+\frac{d^2\mathbf{F}}{d\boldsymbol{\beta}^2}\,(\mathbf{I}_p\otimes(\mathbf{D}^{-1}\mathbf{F})\otimes\ \mathbf{I}_p)$$
$$+\left(\left(\frac{d\mathbf{F}}{d\boldsymbol{\beta}}\,(\mathbf{I}_p\otimes\mathbf{D}^{-1})\right)\otimes\ \mathrm{vec}'\mathbf{I}_p\right)\mathbf{K}_{mp,p^2}$$
$$\left(\mathbf{I}_p\otimes\mathbf{K}_{p,p}\otimes\mathbf{I}_m\right)\left(\mathbf{I}_{p^2}\otimes\left(\frac{d\mathbf{F}}{d\boldsymbol{\beta}}\right)'\right).$$