SLU

Epsilon Open Archive http://epsilon.slu.se

A Method for Designing Nonlinear Univariate Calibration

Johannes Forkman

Department of Energy and Technology
Swedish University of Agricultural Sciences
Box 7032, SE-750 07 Uppsala, Sweden
Johannes.Forkman@vpe.slu.se

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) to minimize the errors in the inverse predictions. The curve parameters are assumed to vary randomly between calibrations, with known expected value and known 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.

Key words: Dose-Response, Four-parameter logistic function, Inverse prediction, Michaelis–Menten function, Nonlinear regression, Optimal design

1   Introduction

Calibration curves are used in measuring instruments for determining physical quantities and in analytical procedures for determining concentrations of analytes in biological samples. Mathematically the calibration curve, established by the calibration, is a function from a physical quantity to a response. The calibration curve is fitted to measurements of calibrators (standards) with known physical quantities. By inverse prediction, responses are translated into estimates of physical quantities through the calibration curve. The problem discussed in this article is how to select the design points (i.e., the physical quantities of the calibrators) in such a way that the inverse predictions are optimized with respect to a design criterion.

Consider, for example, the calibration curve shown in Figure 1a, which illustrates a relationship between response and concentration in a measuring range [2, 200] μg/L. Responses of calibrators with concentrations of 2, 10, 50, 100, and 200 μg/L have been measured. A four-parameter logistic function has been fitted by the method of weighted least squares. The fitted calibration can be used for estimating concentrations of samples with unknown
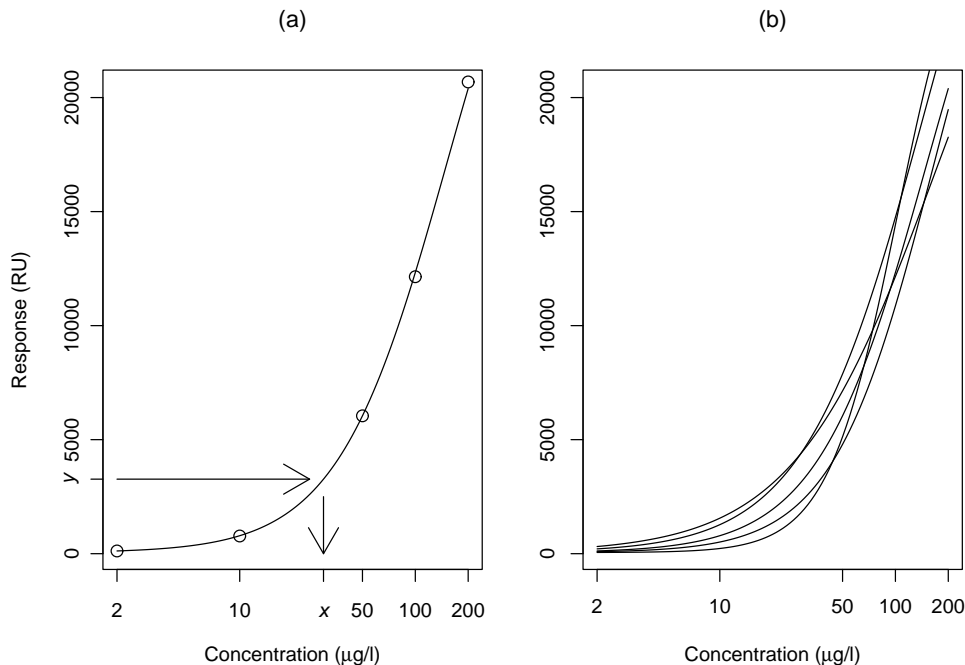
Figure 1: (a) Use of calibration curve for inverse prediction, (b) True relationships between response and concentration, under varying conditions.

concentrations, as indicated by the arrows. In Section 3.2 we investigate whether $\boldsymbol{\xi} = (2, 10, 50, 100, 200)'$ is an optimal design.

New batches of reagents could alter the relationship between response and concentration. Figure 1b illustrates possible true relationships at five days with different batches of reagents. This variation between the curves, caused by changes in reagents or in other influential factors, such as temperature or humidity of the air, could be considered random.

It is a common concern in nonlinear regression that the optimal design is dependent on the curve parameter values, which are not known. Designs are often optimized locally, that is, conditioned on some previously determined fixed curve parameter vector $\boldsymbol{\beta}$. For example, an optimal design could be obtained for one of the curves in Figure 1b, ignoring the random variation between the curves. Locally optimal designs for the Michaelis–Menten model were studied by Dette and Wong (1999) and Boer, Rasch, and Hendrix (2000). A class of models including the three-parameter logistic function was studied by Melas (2004). Gauchi and Pázman (2006)

2

contributed to the theory of locally optimal designs by developing a method for stochastic optimization.

A second approach is to look for a Bayesian optimal design. According to this approach, the curve parameters belong to a prespecified multivariate distribution. The researcher is required to give a full specification of the multivariate distribution for the curve parameters. Bayesian methods could be used not only for design, but also for calibration and inverse prediction by including a prior probability density for the curve parameters and the unknown samples (Osborne 1991). Bayesian optimal designs were reviewed by Chaloner and Verdinelli (1995) and Müller (1999). Clyde and Chaloner (2002) proposed a Bayesian method that could be used to find an optimal design for a given number of design points. Bayesian optimal design for the Michaelis–Menten model has been described by Song and Wong (1998) and Matthews and Allcock (2004).

In this article we suggest a third approach that falls between the two standard approaches. The theory for locally optimal designs, which is reviewed in Section 2.2, is extended in Section 2.3 by assuming that the curve parameters are random with expected value $\boldsymbol{\beta}_0$ and covariance matrix $\boldsymbol{\Sigma}$. The distribution of the curve parameters is not specified, and Bayesian methods are not used for calibration and inverse prediction. Calibrations by the method of least squares are presumed, without assumptions about randomness in the curve parameters. The new design method requires prior specification of the covariance matrix $\boldsymbol{\Sigma}$, with information about how much the curve parameters are likely to vary, (as a result of, e.g., changes in temperature), and to what extent the parameters are correlated. The variances and covariances of $\boldsymbol{\Sigma}$ must be specified based on previous experience of the measuring instrument or the analytical procedure, perhaps obtained in validation studies. Locally optimal designs are obtained as a special case when $\boldsymbol{\Sigma} = \mathbf{0}$. The priors are expressed by means and covariances, as in Bayes linear statistics (Goldstein and Wooff 2007), but the theory is not build on the concept of adjusted expectations.

Dette, and Biedermann (2003), Dette, Melas, and Pepelyshev (2003), Dette, Melas, Pepelyshev, and Strigul (2003, 2005) and Dette, Melas, and Wong (2005) have suggested maximin D- and E-optimal designs for the Michaelis–Menten and Monod models. In the context of generalized linear models, Dror and Steinberg (2006) and Woods, Lewis, Eccleston, and Russell (2006) have proposed methods for constructing robust designs based on the criterion for D-optimality. Mathew and Sinha (2001) have considered D-, E-, and A-optimal designs for logistic regression, and Vila and Gauchi (2007) have proposed designs that minimize exact parameter confidence ellipsoids.

In Section 2.4 it is suggested that the errors in the inverse predictions are minimized over the measuring range. Similar criteria were proposed for linear calibration by Ott and Myers (1968) and Buonaccorsi (1986) and for nonlinear calibration by Rocke and Jones (1997) and François, Govaerts, and Boulanger (2004). For analytical procedures the coefficient of variation, recommended by for example De Silva et al. (2003), and the precision profile (Dudley et al. 1985) are suggested as design criterion.

2 Method Theory

2.1 Model for Calibration

Assume that the expected response is a strictly monotonic continuous function, $f(\xi, \boldsymbol{\beta})$, of the physical quantity ξ, with parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ as in any of the curves in Figure 1b. Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$ be a vector of design points, for example, $\boldsymbol{\xi} = (2, 10, 50, 100, 200)'$. Let $f_i(\boldsymbol{\beta}) = f(\xi_i, \boldsymbol{\beta})$ and $\mathbf{f}(\boldsymbol{\beta}) = (f_1(\boldsymbol{\beta}), f_2(\boldsymbol{\beta}), \ldots, f_m(\boldsymbol{\beta}))'$, so that $\mathbf{f}(\boldsymbol{\beta})$ includes expected responses at the design points. Let $\mathbf{y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\epsilon}$, where the error vector $\boldsymbol{\epsilon}$ has expected value $\mathbf{0}$ and diagonal covariance matrix $\mathbf{D}$. Thus $\mathbf{y}$ is the vector of responses obtained in the calibration, indicated by circles in Figure 1a. Let $\mathbf{F}(\boldsymbol{\beta})$ denote the $m \times p$ matrix of partial derivatives $\partial f_i(\boldsymbol{\beta})/\partial \beta_j$. Assume that the rank of $\mathbf{F}(\boldsymbol{\beta})$ is p and that all required derivatives and moments exist. Let ξ, without index, denote the unknown parameter to 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 weighted least squares estimate of $\boldsymbol{\beta}$ is the $\mathbf{b}$ that minimizes the generalized least squares criterion, so that

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

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. If the number of distinct design points is finite, then $\mathbf{b}$ is consistent as the number of replicates per design point increases (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

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

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 $\mathbf{b} = \boldsymbol{\beta} + (\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{F}(\boldsymbol{\beta}))^{-1}\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\boldsymbol{\epsilon}$. 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, as illustrated in Figure 1a, by $x = f^{-1}(y, \mathbf{b})$, where f^{-1} denotes the inverse of f as a function of ξ. In a neighborhood of $(\mu, \boldsymbol{\beta})$,

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

where δ is the radius of the neighborhood and

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

Note, as is 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

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

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

2.2 Fixed Curve Parameters

In this section we study the variance and expected value of the inverse prediction x of ξ, conditioned on $\boldsymbol{\beta}$. For example, we consider the errors in the inverse prediction given one of the five true relationships between response and concentration shown in Figure 1b. We assume that we know the parameters $\boldsymbol{\beta}$ for this specific curve.

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

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

because y and $\mathbf{b}$ are independent. The first term in (8) is the variance, σ^2,

of the observation in response transformed through linear approximation into the variance in inverse prediction. The second term is the asymptotic covariance matrix, $\mathbf{V}(\boldsymbol{\beta})$, of the curve parameter estimators transformed through linear approximation into the variance in inverse prediction. Thus the approximate conditional variance in x is composed of two parts, one part caused by the variation in the measurement of the sample with unknown ξ and the other 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 estimators (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 $z_i(\boldsymbol{\beta}) = -\mathrm{tr}\{\mathbf{V}(\boldsymbol{\beta})\mathbf{A}_i(\boldsymbol{\beta})\}/2$, $i = 1, 2, \ldots, m$. 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 the bias is approximately equal to $c(\boldsymbol{\beta}) = \mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{F}'(\boldsymbol{\beta})\mathbf{D}^{-1}\mathbf{z}(\boldsymbol{\beta})$.

2.3   Random Curve Parameters

In Section 2.2, we assumed that $\boldsymbol{\beta}$ was a fixed known parameter vector, corresponding to, for example, a single curve in Figure 1b. 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 thus assume that the true relationship can vary randomly, due to, for example, random variation between batches of reagents, as illustrated by the five curves in Figure 1b. 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. Thus, by (8) and (9), the variance in the inverse prediction is

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

where calculation of $k(\boldsymbol{\beta})$ and $\mathbf{g}(\boldsymbol{\beta})$ for a given ξ is made possible by (6) and (7). The last term in (10) is the variance in the bias, which often is 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

expand $k^2(\boldsymbol{\beta})$ 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 $\mathbf{g}'(\boldsymbol{\beta})\mathbf{V}(\boldsymbol{\beta})\mathbf{g}(\boldsymbol{\beta})$. Following Kollo and von Rosen (2005, p. 152), we have

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

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

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

The vec operator stacks the columns of the matrix one under the other, and $\mathbf{A}^{\otimes k}$ is the Kronecker product of k matrices $\mathbf{A}$. Similarly,

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

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

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

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)), can be neglected. The $p \times p$ matrix $d^2(\mathbf{g}'\mathbf{V}\mathbf{g})/d\boldsymbol{\beta}^2$ in (13) can be calculated through the first and second-order derivatives of $\mathbf{g}(\boldsymbol{\beta})$ and $\mathbf{F}(\boldsymbol{\beta})$. The formulas are given in the Appendix.

7

By (9), the unconditional expectation is

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

By 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 $\boldsymbol{\beta} = \boldsymbol{\beta}_0$,

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

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

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

Because $E(\boldsymbol{\beta} - \boldsymbol{\beta}_0) = \mathbf{0}$, and because the last two terms in (14) are often small, we have

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

2.4 Design Criterion

In Section 2.2 and 2.3 we showed 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 m 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 can be the variance or the mean squared error, but it also can be the coefficient of variation γ as defined by

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

If $\boldsymbol{\beta}$ is considered fixed and the search for a locally optimal design is made given a known $\boldsymbol{\beta}$, then (8) and (9) can 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 (13) and (15).

Because ξ is unknown, we want to minimize the function h_ξ for all ξ's that belong to the measuring range. Consequently, the design criterion that

we want to minimize over $\boldsymbol{\xi}$ is

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

where π_ξ is a prior probability density function of ξ. Typically, we assume that ξ belongs to a measuring range $[\alpha, \omega]$. 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 $[\alpha, \omega]$.

Table 1: Notation

$\mathbf{A}_i$	$\partial^2 f_i(\boldsymbol{\beta})/(\partial\beta_r\partial\beta_s)$	x	inverse prediction of ξ
$\mathbf{b}$	GLS estimate of $\boldsymbol{\beta}$	y	sample response
$\mathbf{D}$	covariance matrix of $\mathbf{y}$	$\mathbf{y}$	design point responses
e	$y - \mu$	$\mathbf{z}$	vector of $-\mathrm{tr}\{\mathbf{VA}_i\}/2$
f	calibration curve function	$\boldsymbol{\beta}$	curve parameters
$\mathbf{f}$	$E(\mathbf{y})$	$\boldsymbol{\beta}_0$	$E(\boldsymbol{\beta})$
$\mathbf{F}$	Matrix of $\partial f_i(\boldsymbol{\beta})/\partial\beta_j$	$\boldsymbol{\epsilon}$	$\mathbf{y} - \mathbf{f}$
$\mathbf{g}$	$\partial f^{-1}/\partial\mathbf{b}$	μ	$E(y)$
h_ξ	function to be minimized	ξ	unknown sample value
k	$\partial f^{-1}/\partial y$	$\boldsymbol{\xi}$	vector of m design points
m	number of design points	π_ξ	prior density of ξ
p	number of curve parameters	σ^2	variance in y
$\mathbf{V}$	asymptotic cov. matrix of $\mathbf{b}$	$\boldsymbol{\Sigma}$	covariance matrix of $\boldsymbol{\beta}$

3   Examples

All calculations in the following examples were done using MATLAB 6.5 (Mathworks, Natick, MA). The integrals were calculated by the trapezoidal rule based on 1,000 equally spaced points, on the scale considered.

3.1   The Michaelis–Menten Function

The Michaelis–Menten function,

$$f(\xi, \boldsymbol{\beta}) = \frac{\beta_2\xi}{\beta_1 + \xi}, \tag{18}$$

9

is used in kinetics to describe the reaction rate as a function of substrate concentration ξ. The parameter β_2, often called V_{max}, is the maximum velocity, (i.e. the asymptotic reaction rate when the enzyme becomes saturated with substrate), and β_1 is the Michaelis constant K_m, which is the concentration at which the rate of the enzyme reaction is half-maximal. We consider calibration for enzymatic determination of substrate concentration. Calibration curves described by (18) are shown in Figures 2a and 2b, with the measuring range $[10, 100]$ displayed on a logarithmic scale.

One reasonable design criterion is the average variance over the measuring range in the inverse predictions. The variances can be approximately calculated by (8) for a locally optimal design or by (13) for an optimal design when the curve parameters vary. But only the second term on the right side of (8) and the second and fourth terms of the right side of (13) are dependent on the design. Thus, the average variation in the inverse predictions also is minimized if we use

$$\frac{1}{100 - 10} \int_{10}^{100} \left(\mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}_0) \ + \frac{1}{2} \ \text{vec}'\boldsymbol{\Sigma} \ \text{vec}\left(\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2} \right)' \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \right) d\xi \quad (19)$$

as design criterion, where $\boldsymbol{\Sigma} = \mathbf{0}$ for a locally optimal design.

We look for an integer-valued two-point design that minimizes (19), given that the expected calibration curve is the solid line in Figure 2, with $\boldsymbol{\beta}_0 = (15, 100)'$. We assume that the variance in the measurements of the velocity is constant. Because this variance is a diagonal element of $\mathbf{D}$, it is a multiplicative factor in (19), and we can assume without loss of generality that it equals 1. A full investigation of all $91 \cdot 90/2 = 4,095$ possible designs with two distinct integers in $[10, 100]$ reveals that $\boldsymbol{\xi} = (39, 100)'$ minimizes (19) when $\boldsymbol{\Sigma} = \mathbf{0}$. Thus $(39, 100)'$ is the locally optimal integer-valued design, with design criterion (19). The average variance with this design is 15.373. In comparison, the D-optimal design $(150/13, 100)'$ (Matthews and Allcock 2004) gives an average variance of 16.973.

The slope of the calibration curve is small in the upper part of the measuring range, making the variances in the inverse predictions larger in the upper part than in the lower part. But substrate concentrations may be high only rarely; in this case, it may be sensible to use a design criterion that minimizes the variance on the logarithmic scale. Thus we consider

$$\frac{1}{\log 10} \int_{\log 10}^{\log 100} \left(\mathbf{g}'\mathbf{V}\mathbf{g}(\boldsymbol{\beta}_0) \ + \frac{1}{2} \ \text{vec}'\boldsymbol{\Sigma} \ \text{vec}\left(\frac{d^2(\mathbf{g}'\mathbf{V}\mathbf{g})}{d\boldsymbol{\beta}^2} \right)' \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \right) d(\log \xi)$$
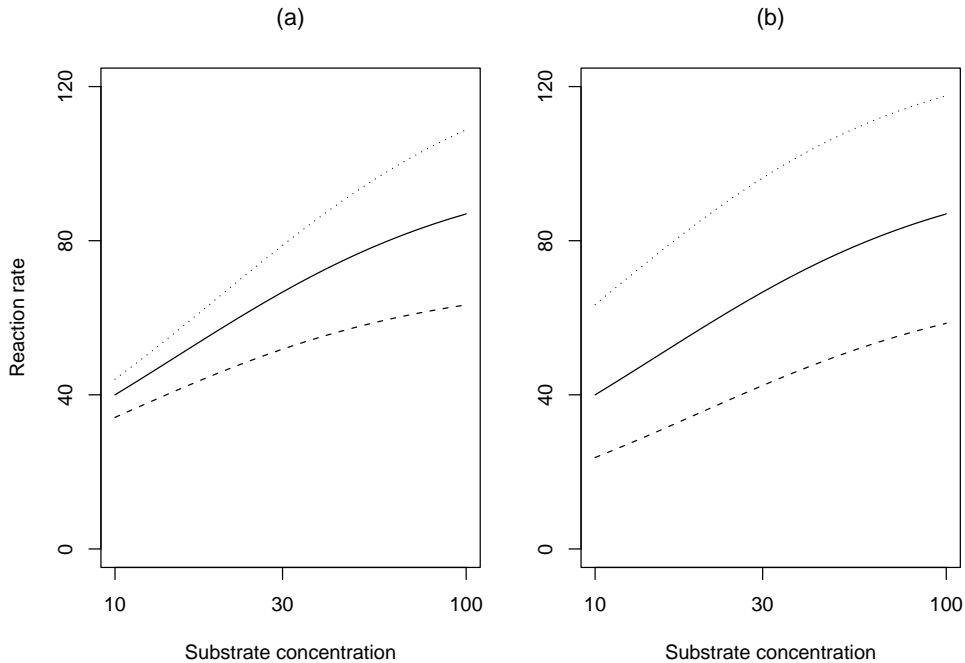
$$(20)$$

Figure 2: Michaelis–Menten functions with $\boldsymbol{\beta} = (15, 100)'$ (solid), (a) $(19.5, 130)'$ (dotted), $(10.5, 70)'$ (dashed), (b) $(10.5, 130)'$ (dotted), $(19.5, 70)'$ (dashed).

as a design criterion. An investigation of all 4,095 possible designs reveals that $(29, 100)'$ minimizes (20) when $\boldsymbol{\Sigma} = \mathbf{0}$, and the minimum is 7.942.

Even if we know from experience that a typical curve has parameters $(15, 100)'$, we may wonder whether the locally optimal design $(29, 100)'$ is optimal when the variation in the parameters also is taken into account. To answer this question, we need to assume that

$$\boldsymbol{\Sigma} = \left(\begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{array} \right),$$

with $\sigma_1 > 0$ and $\sigma_2 > 0$. We may have reason to believe, based on theory or experience, that the standard deviation σ_1 in the Michaelis constant β_1 is approximately 4.5, and that the standard deviation σ_2 in the maximum velocity β_2 is approximately 30. Temperature variation could be the main cause of these variances. Assume that we know that high temperatures usually give high maximum velocities as well as high Michaelis constants, so that

11

β_1 and β_2 are positively correlated. If $\rho = .5$, then a full investigation of all possible integer-valued designs reveals that $(30, 100)'$ is optimal, with design criterion (20) equal to 12.379. Further investigations show that $(30, 100)'$ is optimal also when ρ is 0 or .9, but $(29, 100)'$ is optimal when ρ is $-.5$ or $-.9$. Thus the optimal design does not differ from the locally optimal design if the parameters are negatively correlated ($\rho = .5$ or $\rho = .9$). This could be understood by comparing Figures 2a and 2b. In Figure 2a, with positively correlated parameters, the derivatives change greatly, but in Figure 2b, with negatively correlated parameters, the curves have similar derivatives. It is the changes in the derivatives of the calibration curves that result in changes in the variances in the inverse predictions.

3.2 The Four-Parameter Logistic Function

ImmunoCAP ECP (Phadia AB, Uppsala, Sweden) is an immunoassay method for detecting eosiniphil catonic protein in human sera. The measuring range is $2 - 200$ μg/L. Calibration is made with the four-parameter logistic function

$$f(\xi, \boldsymbol{\beta}) = \beta_2 + \frac{\beta_1 - \beta_2}{1 + (\xi/\beta_3)^{\beta_4}}. \tag{21}$$

The parameter β_1 is the expected response at concentration $\xi = 0$ μg/L, and β_2 is the limit of f when $\xi \to \infty$. When ξ equals β_3, the expected response equals $(\beta_1 + \beta_2)/2$. The slope of the curve is controlled by β_4. By (21), the inverse of the four-parameter logistic function is

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

An experiment can be performed to study the parameters of calibration curves established at varying temperatures and with varying batches of reagents. This study may indicate that the middle curve in Figure 1b, given by the parameter vector $\boldsymbol{\beta}_0 = (40, 34000, 150, 1.4)'$, is typical for ImmunoCAP ECP. The study also may show that different batches of reagents produce different calibration curves, as in Figure 1b. Clearly, the analyst needs to calibrate the method, (i.e., estimate the current parameter values), when new batches of reagents are used. The variation in the curve parame-

12

ters estimated in the study may suggest that

$$\mathbf{\Sigma} = \begin{pmatrix} 100 & -7,680 & -80 & 2.4 \\ -7,680 & 10,240,000 & 12,800 & -900 \\ -80 & 12,800 & 400 & -.64 \\ 2.4 & -900 & -.64 & .16 \end{pmatrix} \tag{22}$$

is a covariance matrix that reasonably describes the variation in the calibration curves. We look for an optimal set of design points for calibration given that $\boldsymbol{\beta}$ is a random vector with expected value $\boldsymbol{\beta}_0 = (40, 34000, 150, 1.4)'$ and covariance matrix (22).

To avoid extrapolation, we decide beforehand that the lowest design point should be 2 and the highest 200 μg/L. Thus, if we want to use only five positions in the instrument for calibrators, then we search for an optimal set of three additional design points in the measuring range $[2, 200]$. Let $\widehat{\gamma}$ be an approximate value of the coefficient of variation (16), defined as the square root of the right side of (13) divided by the right side of (15). As design criterion, we use

$$\int_2^{200} h_\xi \pi_\xi \, d\xi = \frac{1}{\log 100} \int_{\log 2}^{\log 200} \widehat{\gamma} \, d(\log \xi), \tag{23}$$

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, and the area under the precision profile is minimized. It makes sense to display the measuring range on the logarithmic scale because subjects are more evenly distributed on the logarithmic scale than on the original scale and because the logistic function (21) represents a logistic regression of the proportion $(\mu - \beta_1)/(\beta_2 - \beta_1)$ on $\log \xi$. The calibration curve (21) in ImmunoCAP ECP is fitted by the method of weighted least squares. The response variance is a power function of the mean. We let $\sigma^2 = \phi \mu^\theta$, and let the diagonal elements of $\mathbf{D}$ be $\phi(f_i(\boldsymbol{\beta}_0))^\theta$, $i = 1, 2, \ldots, 5$, where $\phi = .00067$ and $\theta = 1.88$.

The function fminsearch in MATLAB makes use of the simplex search method by Nelder and Mead (1965) and algorithms given by Lagarias, Reeds, Wright, and Wright (1998). With starting value $(2.00, 10.0, 50.0, 100, 200)'$, this search algorithm converges after 93 iterations at $\boldsymbol{\xi} = (2.00, 6.66, 18.0, 82.3, 200)'$, where the integral (23) equals 2.8389%. Rounded to $(2.00, 7.00, 18.0, 80.0, 200)'$, the average coefficient of variation (23) is only slightly larger: 2.8392%. Figure 3 provides the precision profile for this design. For comparison, (23) equals 5.7813% with the equidistant design $(2.00, 51.5,$
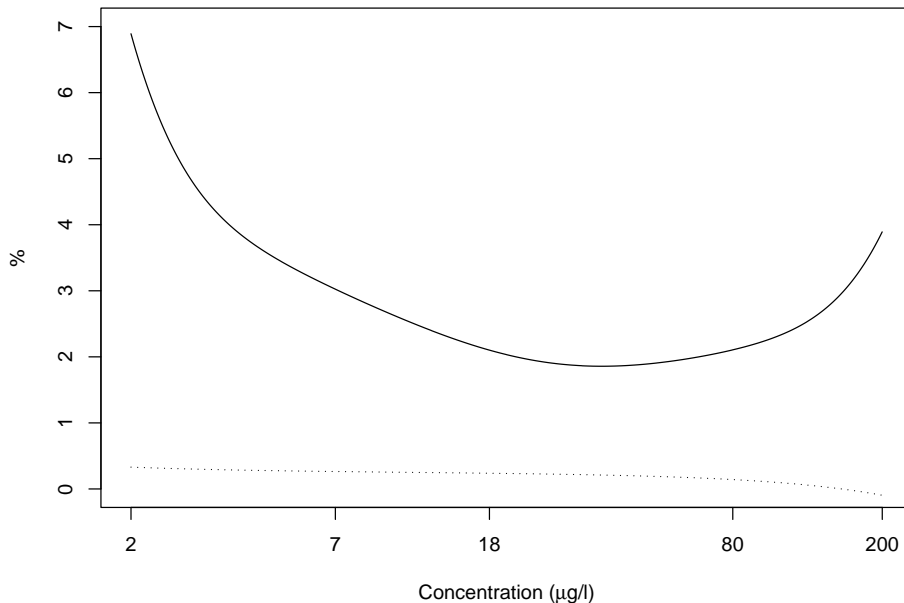
Figure 3: Precision profile for the set of design points $(2.00, 7.00, 18.0, 80.0, 200)'$. The integral of the coefficient of variation (solid line) is minimized over the measuring range on logarithmic scale. The bias (dotted line), illustrated as a percentage of concentration, is small compared with the standard deviation.

101, 150.5, 200)' and equals 2.8582% with the equidistant design on the logarithmic scale $(2.00, 6.32, 20.0, 63.2, 200)'$.

If we instead consider $\boldsymbol{\beta} = (40, 34000, 150, 1.4)'$ to be a fixed parameter vector and calculate $\widehat{\gamma}$ as the square root of the right side of (8), divided by the right side of (9), the algorithm converges after 98 iterations at the locally optimal design $(2.00, 5.70, 13.2, 60.2, 200)'$, where the integral (23) equals 1.9727%. In comparison, if $\widehat{\gamma}$ is calculated by (13) and (15), then the integral (23) equals 2.8512% at the locally optimal design.

4   Method Summary

1. Determine, based on pre-studies and theoretical knowledge, the expected curve parameter values $\boldsymbol{\beta}_0$ and their variances and covariances $\boldsymbol{\Sigma}$. For a locally optimal design, set $\boldsymbol{\Sigma} = \mathbf{0}$.

14

2. Determine, based on pre-studies and theoretical knowledge, the variance σ in response of a sample with concentration ξ. This could be a constant or a function of ξ. Also determine the diagonal matrix $\mathbf{D}$, with the variances of the design points in response. These variances can be equal and constant or a function of the design points.

3. Determine the function h_ξ, which should be as small as possible for every concentration ξ. This function can be, for example, the variance, the mean squared error or the coefficient of variation.

4. Determine weights for the summation of h_ξ over the measuring range by setting the probability density function π_ξ included in the design criterion (17). If all parts of the measuring range $[\alpha, \omega]$ are equally important, then set $\pi_\xi = 1/(\omega - \alpha)$.

5. Differentiate the calibration curve function f with respect to $\boldsymbol{\beta}$ to obtain $\mathbf{F}(\boldsymbol{\beta})$. If the bias is included in the design criterion, then find the second-order derivatives of f with respect to $\boldsymbol{\beta}$ to obtain $A_i(\boldsymbol{\beta})$, $i = 1, 2, \ldots, m$. Differentiate the inverse calibration function f^{-1} with respect to y and $\mathbf{b}$ to obtain $k(\boldsymbol{\beta})$ and $\mathbf{g}(\boldsymbol{\beta})$. For an optimal design based on random parameters, (i.e., if $\boldsymbol{\Sigma} \neq \mathbf{0}$), determine the first-order and second-order derivatives of $\mathbf{g}(\boldsymbol{\beta})$ and $\mathbf{F}(\boldsymbol{\beta})$.

6. For a given design $\boldsymbol{\xi}$, the function h_ξ can be calculated by (13) and (15), and the design criterion (17) can be evaluated numerically. Compare different designs by (17). If necessary, use a search algorithm, (e.g., the simplex method), to find the optimal design.

5 Discussion

A locally optimal design is optimized only for a given fixed parameter vector. This is a problem, because the parameters are likely to change. Otherwise, it would not be necessary to calibrate the instrument. A good design performs well under varying curve parameters. Thus we need to consider not only the expected parameter values, but also the variances and covariances. These can be estimated from a data set of measurements performed under varying conditions. An instrument manufacturer can perform precision studies in countries with different climatic conditions and estimate the variation in calibration curves established with different batches of reagents. It is more difficult to make the assumptions required by Bayesian optimization about

the exact mathematical form of the multivariate distribution of the parameters. Bayesian optimization requires a multidimensional integration over the parameter space or the use of Monte Carlo based methods. This is especially challenging in the calibration context when (17) is used as a design criterion, because this criterion involves additional integration.

Using the method suggested in this article it is possible to investigate the effect of varying parameters. The design often is more sensitive to changes in the expected value $\boldsymbol{\beta}_0$ than to changes in the covariance matrix $\boldsymbol{\Sigma}$. As shown in the example of Section 3.1, the design also is sensitive to the choice of scale in the design criterion (17), that is, the choice of the function π_ξ.

We have assumed that the measuring range and the number of design points are determined in advance. In analytical procedures, the measuring range should begin at the quantitation limit, that is, the concentration at which samples can be quantitatively determined with stated acceptable precision and trueness (Clinical and Laboratory Standards Institute 2004), and end before the errors in the inverse predictions become too large (Gottschalk and Dunn 2005). In practice, the number of design points m often is determined by the design of the measuring instrument. Only m positions may be available for calibrators, because the other positions in the instrument should be used for samples with unknown concentrations. When m is large, the optimal design often consists of less than m distinct design points. It is desirable to have as many design points as possible to minimize the errors in the inverse predictions. We also may wish to have at least $p + 1$ distinct design points, where p is the number of curve parameters, so that the model fit can be checked (Lupinacci and Raghavarao 2000). Designs with different numbers of design points can be compared by the design criterion (17). To find the optimal design for m measurements in v replicates, the diagonal covariance matrix $\mathbf{D}$ should be divided by v. When $\mathbf{D}$ includes variance parameters, as in the example in Section 3.2, these must be determined in advance by, for example, the method proposed by Raab (1981).

Many authors have studied the design of nonlinear regression under the assumption of normally distributed measurements. In this article normality is only needed for the estimation of the bias. As illustrated by the example in Section 3.2, the bias in the inverse prediction often is small compared with the standard deviation. It could thus be acceptable to neglect the bias and build the design criterion on a function of the variance alone, as in the example in Section 3.1. Rocke and Jones (1997) noted that only the second term in the variance (8) is dependent on the design. If the integral of (8) is to be minimized, then it suffices to minimize the integral of the second term of (8). Rocke and Jones (1997) suggested maximizing the integral of

the inverse variance. In this case, it might be better to include the first term of (8).

It is usually assumed (e.g. François et al. 2004) that the inverse calibration function has an analytical form. According to Schwenke and Milliken (1991), a confidence interval based on $\mathbf{V}(\boldsymbol{\beta})$ requires a closed form for the inverse function; however, (7) can be used to differentiate the inverse function when no closed form exists.

A software package for symbolic mathematics facilitates calculation of derivatives. Because the design criterion (17) is not necessarily a convex function of the design points, search algorithms may stop at local minima. It could be wise to begin with a grid search over the design space.

Acknowledgments

Appendix: Matrix Derivatives

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

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

so that $\mathbf{F}(\boldsymbol{\beta}) = (d\mathbf{f}/d\boldsymbol{\beta})'$. 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 1, and all of the other elements in that block equal 0. By the rules of matrix differentiation,

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

17

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

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

References

Boer, E. P. J., Rasch, D. A. M. K., and Hendrix, E. M. T. (2000), "Locally Optimal Designs in Non-Linear Regression: A Case Study of the Michaelis–Menten Function," in *Advances in Stochastic Simulation Methods*, eds. N. Balakrishnan, V. B. Melas, and S. Ermakov, Boston: Birkhäuser, pp. 177-188.

Box, M. J. (1971), "Bias in Nonlinear Estimation," *Journal of the Royal Statistical Society*, Ser. B, 32, 171-190.

Buonaccorsi, J. P. (1986), "Design Considerations for Calibration," *Technometrics*, 28, 149-155.

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

Clinical and Laboratory Standards Institute (2004), *Limits of Detection and Quantitation*, document EP17- A, Wayne, PA: D. W. Tholen.

Clyde, M., and Chaloner, K. (2002), "Constrained Design Starategies for Improving Normal Approximations in Nonlinear Regression Problems," *Journal of Statistical Planning and Inference*, 104, 175-196.

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.

Dette, H., and Biedermann, S. (2003), "Robust and Efficient Designs for the Michaelis–Menten Model," *Journal of the American Statistical Association*, 98, 679-686.

Dette, H., Melas, V. B., and Pepelyshev, A. (2003), "Standardized Maximin *E*-optimal Designs for the Michaelis–Menten Model," *Statistica Sinica*, 13, 1147-1163.

Dette, H., Melas, V. B., Pepelyshev, A., and Strigul, N. (2003), "Efficient Design of Experiments in the Monod Model," *Journal of the Royal Statistical Society*, Ser. B, 65, 725-742.

Dette, H., Melas, V. B., Pepelyshev, A., and Strigul, N. (2005), "Robust and Efficient Design of Experiments for the Monod Model," *Journal of Theoretical Biology*, 234, 537-550.

Dette, H., Melas, V. B., and Wong, K. W. (2005), "Optimal Design for Goodness of Fit of the Michaelis–Menten Enzyme Kinetic Function," *Journal of the American Statistical Association*, 100, 1370-1381.

Dette, H., and Wong, W. K. (1999), "E-optimal Designs for the Michaelis–Menten Model," *Statistics & Probability Letters*, 44, 405-408.

Dror, H. A., and Steinberg, D. M. (2006), "Robust Experimental Design for Multivariate Generalized Linear Models," *Technometrics*, 48, 520-529.

Dudley, 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.

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.

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.

Gauchi, J. -P., and Pázman, A. (2006), "Designs in Nonlinear Regression by Stochastic Minimization of Functionals of the Mean Square Error Matrix" *Journal of Statistical Planning and Inference*, 136, 1135-1152.

Goldstein, M., and Wooff, D. (2007), *Bayes Linear Statistics: Theory and Methods*, Chichester, U.K.: Wiley.

Gottschalk, P. G., and Dunn, J. R. (2005), "Determining the Error of Dose Estimates and Minimum and Maximum Acceptable Concentrations from Assays with Nonlinear Dose-Response Curves," *Computer Methods and Programs in Biomedicine*, 80, 204-215.

Jennrich, R. I. (1969), "Asymptotic Properties of Non-Linear Least Squares Estimators," *The Annals of Mathematical Statistics*, 40, 633-643.

Kollo, T., and von Rosen, D. (2005), *Advanced Multivariate Statistics With Matrices*, Dordrecht: Springer.

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.

Lupinacci, P. J., and Raghavarao, D. (2000), "Designs for Testing Lack of Fit for a Nonlinear Dose-Response Curve Model," *Journal of Biopharmaceutical Statistics*, 10, 45-53.

Malinvaud, E. (1970), "The Consistency of Nonlinear Regressions," *The Annals of Mathematical Statistics*, 41, 956-969.

Mathew, T., and Sinha, B. K. (2001), "Optimal Designs for Binary Data Under Logistic Regression," *Journal of Statistical Planning and Inference*, 93, 295-307.

Matthews, J. N. S., and Allcock, G. C. (2004), "Optimal Designs for Michaelis–Menten Kinetic Studies," *Statistics in Medicine*, 23, 477-491.

Melas, V. B (2004), "On the Functional Approach to Optimal Designs for Nonlinear Models," *Journal of Statistical Planning and Inference*, 132, 93-116.

Müller, P. (1999), "Simulation-Based Optimal Design," *Bayesian Statistics*, 6, 459-474.

Nelder, J. A., and Mead, R. (1965), "A Simplex Method for Function Minimization," *Computer Journal*, 7, 308-313.

Osborne, C. (1991), "Statistical Calibration: A Review," *International Statistical Review*, 59, 309-336.

Ott, R. L., and Myers, R. H. (1968), "Optimal Experimental Designs for Estimating the Independent Variable in Regression," *Technometrics*, 10, 811-823.

Raab, G. M. (1981), "Estimation of a Variance Function, With Application to Immunoassay," *Applied Statistics*, 30, 32-40.

Rocke, D. M., and Jones, G. (1997), "Optimal Design for ELISA and Other Forms of Immunoassay," *Technometrics*, 39, 162-170.

Schwenke, J. R., and Milliken, G. A. (1991), "On the Calibration Problem Extended to Nonlinear Models," *Biometrics*, 47, 563-574.

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

Song, D., and Wong, W. K. (1998), "Optimal Two-Point Designs for the Michaelis–Menten Model with Heteroscedastic Errors," *Communications in Statistics, Part A – Theory and Methods*, 27, 1503-1516.

Vila, J. P., and Gauchi, J. -P. (2007), "Optimal Designs Based on Exact Confidence Regions for Parameter Estimation of a Nonlinear Regression Model," *Journal of Statistical Planning and Inference*, 137, 2935-2953.

Woods, D. C., Lewis, S. M., Eccleston, J. A., and Russell, K. G. (2006), "Designs for Generalized Linear Models With Several Variables and Model Uncertainty," *Technometrics*, 48, 284-292.