Implementation of Robust Designs for Straight Line Regression

by

Md. SHAKHAWAT HOSSAIN

Master of Science Project
December 2002

Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton, Alberta
Canada
ACKNOWLEDGEMENTS

First of all, I would like to express my deepest sense of gratitude to my supervisor Douglas P. Wiens for his guidance, encouragement and excellent advice on my project and throughout my course work. Without his support, financial and otherwise, this project would have been impossible.

Also, my sincere thanks to my friends A. Adewale, M. Shofiqul Islam, Xiaojian Xu and a whole lot of others whose assistance and encouragement have seen me to the completion of my program.
1 Introduction

Consider the construction of a design for the straight line regression model, robust against unspecified contamination of the response function. Suppose that the experimenter fits, by least squares, a regression model

\[ y = \theta_0 + \theta_1 x + \varepsilon \]

\[ E(y|x) = \theta_0 + \theta_1 x. \]

Let \( y_i, i = 1, 2, 3, \ldots, n \), denote the independent observations taken at \( x_i, i = 1, 2, 3, \ldots, n \), then the simple linear regression for these data can be written as

\[ y_i = \theta_0 + \theta_1 x_i + \varepsilon_i \quad (1) \]

\[ E(y|x_i) = \theta_0 + \theta_1 x_i. \quad (2) \]

The errors terms \( \varepsilon_i \) are uncorrelated with mean zero and common variance \( \sigma^2 \). The regressor \( x_i \) varies over a design space \( S \). The experimenter is concerned that the true model might be approximated by (1), with a more precise description being

\[ E(y|x_i) = \theta_0 + \theta_1 x_i + f(x_i), \quad (3) \]

where the function \( f(x) \) is unknown but ‘small’. Due to the unknown nature of the true model (3), model (2) is usually fitted after observing \( n \) data points \( \{(y_i, x_i)\}_{i=1}^{n} \). The least squares estimate \( \hat{\theta} \) of \( \theta \) and \( \hat{Y} = z^T(x)\hat{\theta} \) (where \( z^T(x) = (1 \ x) \)) of \( E(y|x) \) are possibly biased because of unspecified contamination of the response function. In this situation the experimenter will choose the design points that yield good estimates \( \hat{\theta} \) and \( \hat{Y} \) as well as offering some protection against the contamination.
The optimal placement of design points was found by Huber (1975, 1981) under the assumption that the contamination function $f(x)$ belongs to

$$F_2 = \left\{ f : \int_S f^2(x) \, dx \leq \eta^2, \int_S f(x) \, dx = \int_S xf(x) \, dx = 0 \right\}.$$ 

The radius $\eta$ is assumed known. The first condition is required in order that the contamination function $f(x)$ be small, so the linear part of (3) is still the dominant term and the errors due to bias do not swamp those due to variance. The second condition is to guarantee that the true $\theta$ is uniquely defined and is made without loss of generality. To see this, suppose a linear function $z^T(x)\theta_0$ of the regressors is the best approximation of the experimental outcome, $E(y|x)$, in the sense that

$$\theta_0 = \arg\min_\theta \left\{ \int_S \left\{ E[y|x] - z^T(x)\theta \right\}^2 \, dx \right\}.$$ 

Define $f(x) = E[y|x] - z^T(x)\theta_0$, then we have the model (3) and $f(x)$ satisfies

$$\int_S z(x)f(x) \, dx = 0.$$ 

If $\int_S z(x)z^T(x) \, dx$ is positive definite then the parameter $\theta_0$ is unique. Now consider the design measure $\xi$ for a straight line regression over $S$. The optimal design measure $\xi_0$ has the symmetric density $m_0(x)$ (p. 247, Huber [1981]). So it is not implementable yet. Therefore, practically we have to approximate it using a discrete design measure $\xi_n$. The following are some of the methods by which this can be done.

1. One could randomly choose the $x_i$ from the density $m_0(x)$. Equivalently let $u_1, u_2, u_3, \ldots, u_n$ be generated values of uniform random variables on $(0,1)$ and put $x_i = \xi_0^{-1}(u_i)$.

To prove that $\xi_n \xrightarrow{L} \xi_0$, first we show that $\xi_0^{-1}(U) \leq x \iff U \leq \xi_0(x)$. Suppose

$$x \geq \xi_0^{-1}(U),$$
then
\[ \xi_0 (x) \geq \xi_0 [\xi_0^{-1}(U)] \geq U \]
since \( \xi_0 (x) \) is a nondecreasing function. Thus
\[ \xi_0 (x) \geq U. \]
Conversely, suppose that
\[ U \leq \xi_0 (x), \]
then
\[ \xi_0^{-1}(U) \leq \xi_0^{-1}[\xi_0 (x)] \leq x, \]
since \( \xi_0^{-1}(x) \) is a nondecreasing function. Thus
\[ \xi_0^{-1}(U) \leq x. \]
Hence
\[ \xi_0^{-1}(U) \leq x \iff U \leq \xi_0 (x), \]
implying that
\[ P(\xi_0^{-1}(U) \leq x) = P(U \leq \xi_0 (x)). \]
So if \( X = \xi_0^{-1}(U) \) represents the r.v. chosen in this manner, and \( P_n \) denotes the empirical probability, then
\[ P_n (X \leq x) = P_n (\xi_0^{-1}(U) \leq x) = P_n (U \leq \xi_0 (x)) \rightarrow \xi_0 (x). \]
Hence
\[ \xi_n \xrightarrow{L} \xi_0. \]

2. A more systematic approach is to put the observations at quantiles of the optimal design. If we choose \( x_i = \xi_0^{-1}\left(\frac{i-1}{n-1}\right) \) then the empirical distribution of the design tends weakly to the true optimal design [3].
3. Another way to enforce convergence of the empirical design is by matching up the moments. Carleman’s condition implies that if \( E_{\xi_n}(x^k) \rightarrow E_{\xi_0}(x^k) \) for \( k = 1, 2, 3, \ldots \) then \( \xi_n \xrightarrow{L} \xi_0 \). We choose \( x_1, x_2, x_3, \ldots, x_n \) symmetrically to solve

\[
\frac{1}{n} \sum x_i^{2k} = E_{\xi_0}(X^{2k})
\]

for as many integral values of \( k \) as possible.

We will measure the quality of each method through the MSE (Mean Square Error)

\[
\text{MSE}(x) = E \left[ \left( z^T(x) \hat{\theta} - E [y|x] \right)^2 \right]
\]

where \( z^T(x) = (1, x) \), \( \hat{\theta} = (\hat{\theta}_0, \hat{\theta}_1)^T \). This MSE includes the variance of \( z^T(x)\hat{\theta} \) around its mean and also bias.

We consider the overall loss function (integrated mean square error)

\[
\text{IMSE} = \int_S \text{MSE}(x) \, dx = \int_S E \left[ \left( z^T(x) \hat{\theta} - E [y|x] \right)^2 \right] \, dx.
\]

Of the above three methods, we are evaluating \( MSE(f, \xi) \) for a particular \( f(x) \) and then at \( \xi_n \), where \( \xi_n \) is attained by each of the three above methods. We will then determine which method results in the smallest values of \( MSE(f, \xi) \).

### 2 Classical Regression Design

In simple linear regression we have the data \( \{(y_i, x_i)\}_{i=1}^n \) and we wish to describe the association between \( x \) and \( y \) by a statistical model. Consider an example:
Suppose the experimenter conducted an experiment to measure the water purity \((y)\) which is a function of chlorine \((x)\). We assume that \(x\) is coded such that \(x \in [-1, 1]\). When observations (water purity) are subject to experimental error, the \(n\) observations are given by

\[
y_i = \theta_0 + \theta_1 x_i + \epsilon_i \quad i = 1, 2, 3, ..., n
\]

(4)

where \(x_i\) is the amount of chlorine added and \(y_i\) is the level of water purity corresponding to \(x_i\). We assume that the error terms are uncorrelated with mean zero and variance \(\sigma^2\). The parameters \(\theta_0\) and \(\theta_1\) are unknown regression coefficients. We use the least squares estimates \(\hat{\theta} = (X^T X)^{-1} X^T y\), where

\[
X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & x_4 & \cdots & x_n \end{pmatrix}
\]

The variances of the component estimates are

\[
n VAR(\hat{\theta}_1) = \frac{\sigma^2}{S_x^2} \quad \text{and} \quad n VAR(\hat{\theta}_0) = \sigma^2 \left( 1 + \frac{\bar{x}^2}{S_x^2} \right)
\]

where

\[
S_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2.
\]

Our objective here is to choose \(x_i\)'s which will improve the quality of our estimates \(\theta_0\) and \(\theta_1\). One way to improve the quality is to put half of \(x_i\)'s at the lowest \(x\)-values and the other at the highest \(x\)-values. In that case, \(\bar{x} = 0\) gives the minimum variances and \(S_x^2 = \text{maximum}\). This minimizes both \(VAR(\hat{\theta}_1)\) and \(VAR(\hat{\theta}_0)\). Since the model (4) is not exactly correct, there is some bias in the model. That's why we consider another method that will take care against bias.
3 Arbitrary IMSE Design

Consider the design measure \( \xi(x) = \text{fraction of observations made at } x \) which we hope results in small values of IMSE (Integrated Mean Square Error).

Now

\[
\text{MSE}(x) = E \left[ \{ z^T(x) \hat{\theta} - z^T(x)\theta - f(x) \}^2 \right]
\]

\[
= z^T(x) E \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] z(x) - 2f(x) z^T(x) E(\hat{\theta} - \theta) + f^2(x)
\]

Since \( \int_S z(x) f(x) \, dx = 0 \), we get

\[
\text{IMSE} = \int_S \text{MSE}(x) \, dx
\]

\[
= \int_S z^T(x) \text{MSE} z(x) \, dx + \int_S f^2(x) \, dx \quad (5)
\]

where \( \text{MSE} = E \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \) is the MSE matrix of \( \hat{\theta} \). Now

\[
\text{MSE} = E \left[ \left( \hat{\theta} - E(\hat{\theta}) \right) \left( \hat{\theta} - E(\hat{\theta}) \right)^T \right] + \left( E(\hat{\theta}) - \theta \right) \left( E(\hat{\theta}) - \theta \right)^T
\]

\[
= \text{Cov}(\hat{\theta}) + \text{bias} \cdot \text{bias}^T.
\]

In terms of design measure \( \xi \) (consider it as the empirical distribution function of \( \{x_1, x_2, x_3, ..., x_n\} \), placing mass \( n^{-1} \) at each \( x_i \)), the covariance matrix is

\[
\text{Cov}(\hat{\theta}) = \sigma^2 (X^T X)^{-1}
\]

\[
= \sigma^2 \left( \sum z(x_i) z^T(x_i) \right)^{-1}
\]

\[
= \frac{\sigma^2}{n} \left( \frac{1}{n} \sum z(x_i) z^T(x_i) \right)^{-1}
\]

\[
= \frac{\sigma^2}{n} \left\{ E(\xi) [z(x) z^T(x)] \right\}^{-1}
\]

\[
= \frac{\sigma^2}{n} \text{B}^{-1}(\xi)
\]
where \( B(\xi) = E_\xi(z(x) z^T(x)) \) depends on the design \( \xi \), not on \( f \).

The bias is

\[
\text{bias} = E(\hat{\theta}) - \theta = E\left( (X^T X)^{-1} X^T y \right) - \theta = n(X^T X)^{-1} \cdot \frac{1}{n} \sum z(x_i) f(x_i) = n [nB(\xi)]^{-1} E_\xi[z(x) f(x)] = B^{-1}(\xi)b(f, \xi),
\]

where

\[
b(f, \xi) = E_\xi[z(x) f(x)] = E_\xi\left[ \left( \frac{1}{n} \sum f(x_i) \right) \right] = \frac{1}{n} \sum x_i f(x_i). \tag{6}
\]

Again

\[
A = \int_S z(x) z^T(x) \, dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left( \frac{1}{x} \right) (1 \, x) \, dx = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12} \end{pmatrix}. \tag{7}
\]

So from (5)

\[
IMSE = \int_S \text{tr} \left[ z^T(x) \text{MSE} \, z(x) \right] \, dx + \int_S f^2(x) \, dx = \int_S \text{tr} \left[ \text{MSE} \cdot z(x) z^T(x) \right] \, dx + \int_S f^2(x) \, dx = \text{tr} \left[ \text{MSE} \cdot A \right] + \int_S f^2(x) \, dx. \tag{8}
\]
Now
\[
tr(MSE \cdot A) = tr \left[ Cov(\hat{\theta}) \cdot A \right] + bias^T \cdot A \cdot bias
\]
\[
= \frac{\sigma^2}{n} tr \left[ B^{-1}(\xi)A \right]
+ b^T(f, \xi) \cdot B^{-1}(\xi) \cdot A \cdot B^{-1}(\xi) \cdot b(f, \xi).
\] (9)

With
\[
H = B^{-1} AB^{-1},
\]
(9) becomes
\[
tr(MSE \cdot A) = \frac{\sigma^2}{n} tr \left[ B^{-1} A \right]
+ b^T H^{-1} b
\]
\[
= \nu \eta^2 tr \left[ B^{-1}(\xi)A \right]
+ b^T H^{-1} b,
\] (10)

where \( \nu = \frac{\sigma^2}{n \eta^2} \) is the relative importance of variance versus bias.

From (8) and (10) we have
\[
\text{loss} = \eta^2 \left[ \nu tr \left[ B^{-1} A \right] + \frac{b^T H^{-1} b}{\eta^2} + \frac{\int_S f^2(x) dx}{\eta^2} \right]
\] (11)

where \( \int_S f^2(x) dx = \eta^2 \). Then
\[
\int_S \left( \frac{f(x)}{\eta} \right)^2 dx = 1.
\] (12)

Thus
\[
\frac{b^T H^{-1} b}{\eta^2} = \frac{b^T(f, \xi)}{\eta} \frac{H^{-1} b(f, \xi)}{\eta},
\]
and
\[
\frac{b(f, \xi)}{\eta} = b \left( \frac{f}{\eta}, \xi \right),
\]
because
\[
b \left( \frac{f}{\eta}, \xi \right) = E_\xi \left[ z(x), \frac{f}{\eta} \right] = \frac{1}{\eta} E_\xi [z(x), f(x)] = \frac{1}{\eta} b (f, \xi).
\]
Thus by (12) we can assume that \( f(x) \) is normalized so that \( \int_S f^2(x) \, dx = 1 \) and we write \( \frac{\text{loss}}{\eta} \) as \( l(\xi) \). Finally, from (11) we get

\[
  l(\xi) = 1 + \nu tr \left[ B^{-1} A \right] + b^T H^{-1} b
\]  

(13)

where

\[
  b(f, \xi) = \int_S z(x) \, f(x) \, d\xi(x) = \frac{1}{n} \sum z(x_i) f(x_i)
\]

and \( f(x) \) is assumed to satisfy \( \int_S f^2(x) \, dx = 1 \). Since

\[
  B = E_\xi (zz^T) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{n} \sum x_i^2 \end{pmatrix}
\]

(here the off-diagonal elements are zero because our design is symmetric i.e. \( \frac{1}{n} \sum x_i = 0 \)) and using (7) we have

\[
  H^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12} \left( \frac{n}{\sum x_i^2} \right)^2 \end{pmatrix}
\]

Finally using (6), we get

\[
  b^T H^{-1} b = \left( \frac{1}{n} \sum f(x_i) \right)^2 + \frac{1}{12} \left( \sum x_i f(x_i) \right)^2 \cdot \left( \frac{1}{\sum x_i^2} \right)^2 = \tau_n.
\]

From (13), the loss function is

\[
  \ell(\xi) = 1 + \nu \left( 1 + \frac{1}{12 \gamma_n} \right) + \tau_n.
\]

\[ (14) \]

4 Huber’s minimax design for simple linear regression

Huber obtains the minimax design by fitting the straight line regression model although the true response is only approximately linear. We recall
the fitted model and the true model,

\[
\text{Fitted Model : } \quad E(y|x) = \theta_0 + \theta_1 x \quad \text{and} \quad E(y|x_i) = \theta_0 + \theta_1 x_i + f(x).
\]

By minimizing the loss function of Q-optimality (minimize the maximum of the integrated (over \(S\)) MSE of the estimated response \(\hat{Y} = \mathbf{z}^T(x)\hat{\theta}\) over \(F_2\), Huber(1975) obtained robust optimal designs with \(\mathbf{z}^T(x) = (1, x)\) and \(S = \left[\frac{-1}{2}, \frac{1}{2}\right]\). For loss function \(\ell(\xi)\) in (14) the minimax design was discussed in Huber (1975, 1981) and the minimax density function is given by

\[
m_0(x) = (ax^2 + b)^+
\]

where \((\cdot)^+\) is defined in (22). The constants \(a\) and \(b\) depend on the parameter \(\nu = \frac{\sigma^2}{n}\) and they can be determined from the following conditions,

\[
\int_S m_0(x) \, dx = 1 \quad (16)
\]

and

\[
\int_S x^2 m_0(x) \, dx = \gamma, \quad (17)
\]

where \(\gamma\) is then chosen to minimize the maximum loss for fixed \(\nu\).

4.1 Case 1: \(\frac{1}{12} \leq \gamma \leq \frac{3}{20}\).

When \(\frac{1}{12} \leq \gamma \leq \frac{3}{20}\), the optimal design measure \(\xi_0\) has the density described as follows. Using (16) and (17), Huber (1981) obtained the constants \(a\) and \(b\):

\[
a = 15(12\gamma - 1) > 0
\]

and

\[
b = 1 - \frac{5}{4}(12\gamma - 1) > 0.
\]
Then
\[ m_0(x) = 1 + \frac{5}{4}(12\gamma - 1)(12x^2 - 1) \] (18)
and
\[ f_0(x) = \sqrt{\frac{5}{4}}(12x^2 - 1). \] (19)
The parameters \( \nu \) and \( \gamma \) are related by
\[ \nu = 360\gamma^2 (12\gamma - 1). \] (20)
Since \( \gamma \in \left[\frac{1}{12}, \frac{3}{20}\right] \), from (20) we find that \( \nu \in [0, 6.48] \).

4.2 Case 2: \( \frac{3}{20} \leq \gamma \leq \frac{1}{4} \).

Define \( c \in [0, 1) \) by \( c^2 = -\frac{4b}{a} \) then
\[ m_0(x) = \frac{a}{4}(4x^2 - c^2)^+ \] (21)

where
\[ a = \frac{12}{(1-c)^2(1+2c)} \]
and
\[ \gamma = \frac{3 + 6c + 4c^2 + 2c^3}{20(1+2c)}. \]
From (21), we can define the RHS in the following way,
\[ \frac{a}{4}(4x^2 - c^2)^+ = \begin{cases} \frac{a}{4}(4x^2 - c^2), & \text{if } |x| > \frac{c}{2}, \\ 0, & \text{otherwise}. \end{cases} \] (22)
The least favourable function \( f_0(x) \) is
\[ f_0(x) = [m_0(x) - 1] \epsilon \] (23)
where
\[ \epsilon^2 = \frac{5(1-c)(1+2c)^2}{4(1+3c+6c^2+5c^3)}. \] (24)
Finally \(c\) and \(\nu\) are related by

\[
\nu = \frac{18(3 + 6c + 4c^2 + 2c^3)^2}{25(1 - c)^2(1 + 2c)^3}.
\] (25)

Thus for \(c \in [0, 1)\) we find that \(\nu \in [6.48, \infty)\).

The parameter \(\nu\) can be interpreted as representing the relative importance of variance versus bias, in the mind of the experimenter. If \(\nu \to 0\), the bias term completely dominates the problem, and \(\xi_0(x)\) tends to the uniform continuous design which has the density \(m_0(x) \equiv 1\) on \(S\). If \(\nu \to \infty\) then the bias term disappears from the model (13), it becomes pure variance problem and \(\xi_0(x)\) tends to the classical optimal design which takes half of the observations on each of the extreme point of \(S\). See Figure 1.

The loss function is defined by the equation (2.33) in Huber’s book [1], (p. 248)

\[
\text{loss} = 1 + \int_S (m - 1)^2 dx + \nu \left(1 + \frac{1}{12\gamma}\right)
\] (26)
Table 4.1. Loss vs. $\nu$ using minimax loss as at (26)

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>loss</th>
<th>$\nu$</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>1.22</td>
<td>7.07</td>
<td>8.90</td>
</tr>
<tr>
<td>0.48</td>
<td>1.51</td>
<td>8.94</td>
<td>10.87</td>
</tr>
<tr>
<td>0.81</td>
<td>1.88</td>
<td>10.56</td>
<td>12.60</td>
</tr>
<tr>
<td>1.21</td>
<td>2.33</td>
<td>12.88</td>
<td>15.08</td>
</tr>
<tr>
<td>1.69</td>
<td>2.87</td>
<td>16.26</td>
<td>18.70</td>
</tr>
<tr>
<td>2.25</td>
<td>3.51</td>
<td>21.35</td>
<td>24.15</td>
</tr>
<tr>
<td>2.90</td>
<td>4.25</td>
<td>29.39</td>
<td>32.71</td>
</tr>
<tr>
<td>3.64</td>
<td>5.10</td>
<td>42.93</td>
<td>47.05</td>
</tr>
<tr>
<td>4.48</td>
<td>6.06</td>
<td>68.08</td>
<td>73.54</td>
</tr>
<tr>
<td>5.42</td>
<td>7.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.48</td>
<td>8.36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where $m(x)$, $\nu$ and $\gamma$ are defined in (18) and (20) for method 1 and in (21) and (25) for method 2. Using this loss function, Table 4.1 gives numerical results of the loss function for different values of $\nu$

5 Method 1

As discussed in the introduction, in method 1 we solve the equation

$$\xi_0(x) = u$$ (27)

for a random sample $u_1, u_2, u_3, \ldots, u_n$ from uniform(0, 1).

To enforce symmetry we will instead solve (27) for $u_1, u_2, u_3, \ldots, u_{\lfloor n/2 \rfloor}$.

Our design is then $\left\{ \pm x_1, \pm x_2, \pm x_3, \ldots, \pm x_{\lfloor n/2 \rfloor} \right\}$, with “0” appended if $n$ is odd.

The distribution function when $\gamma \in \left[ \frac{1}{12}, \frac{3}{20} \right]$ is

$$\xi_0(x) = \int_{-\frac{1}{2}}^{x} m_0(t) dt$$
which becomes

\[ \xi_0(x) = x + 0.5 + \frac{5}{4}(12\gamma - 1)(4x^3 - x), \]  

while if \( \gamma \in \left[ \frac{3}{20}, \frac{1}{4} \right] \) it is

\[ \xi_0(x) = \begin{cases} 
\frac{a}{24}(1 - 6c^2x - 3c^2 + 8x^3) & \text{if } \frac{1}{2} \leq x \leq \frac{c}{2}, \\
\frac{1}{2}, & \text{if } \frac{c}{2} \leq x \leq \frac{c}{2}, \\
1 - \frac{a}{24}(-8x^3 + 6c^2x + 1 - 3c^2), & \text{if } \frac{c}{2} \leq x \leq \frac{1}{2}, 
\end{cases} \]  

This measure will minimize the right hand side of (13).

Having solved the equations (27) by S-PLUS function “UNIROOT”, we get 25 design points which lie within the interval \( \left[ \frac{-1}{2}, \frac{1}{2} \right] \). To calculate the loss function which is defined in (14) we need to compute \( f_0(x), \tau_n \) and \( \gamma_n \) for those design points. These values of the loss function are random, so we compute this loss 200 times and then average them. Since the loss function \( l(\xi) \) depends on \( \nu \), Table 5.1 gives numerical results of the loss function for different values of \( \nu \), where \( \nu \in [0, 6.48] \) and \( \nu \in [6.48, \infty) \).

Table 5.2 represents the interval that containing 95% values of the loss and the standard deviation of the average loss.

Consider another \( f(x) \) which is cubic function of \( x \),

\[ f(x) = \left( 20x^3 - 3x \right) \sqrt{7}. \]  

This function also satisfies the conditions

\[ \int_S f(x) \, dx = 0 \text{ and } \int_S xf(x) \, dx = 0. \]

In the same way as in method 1 we can evaluate the loss function which is defined in (14). Table 5.3 gives numerical results of the loss function for different values of \( \nu \), again using \( n = 25 \).

Table 5.4 represents the interval that containing 95% values of the loss and standard deviation of the average loss.
### Table 5.1. Loss vs. \( \nu; \) Method 1, \( f_0(x) \)

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>loss</th>
<th>( \nu )</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.08</td>
<td>6.67</td>
<td>12.33</td>
</tr>
<tr>
<td>0.21</td>
<td>1.52</td>
<td>6.71</td>
<td>12.43</td>
</tr>
<tr>
<td>0.48</td>
<td>2.03</td>
<td>6.90</td>
<td>12.83</td>
</tr>
<tr>
<td>0.81</td>
<td>2.66</td>
<td>7.07</td>
<td>13.17</td>
</tr>
<tr>
<td>1.21</td>
<td>3.39</td>
<td>7.82</td>
<td>14.55</td>
</tr>
<tr>
<td>1.69</td>
<td>4.24</td>
<td>8.94</td>
<td>16.45</td>
</tr>
<tr>
<td>2.25</td>
<td>5.20</td>
<td>10.56</td>
<td>19.05</td>
</tr>
<tr>
<td>2.90</td>
<td>6.31</td>
<td>12.88</td>
<td>22.65</td>
</tr>
<tr>
<td>3.64</td>
<td>7.53</td>
<td>16.26</td>
<td>27.75</td>
</tr>
<tr>
<td>4.48</td>
<td>8.89</td>
<td>21.35</td>
<td>35.32</td>
</tr>
<tr>
<td>5.42</td>
<td>10.35</td>
<td>29.39</td>
<td>47.07</td>
</tr>
<tr>
<td>6.48</td>
<td>12.01</td>
<td>42.93</td>
<td>66.60</td>
</tr>
<tr>
<td>6.60</td>
<td>12.18</td>
<td>68.08</td>
<td>102.39</td>
</tr>
<tr>
<td>6.63</td>
<td>12.25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.2. Interval of the loss (95%) and standard deviation

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>interval (95%)</th>
<th>Std.dev.</th>
<th>( \nu )</th>
<th>interval (95%)</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(1.000, 1.341)</td>
<td>0.089</td>
<td>6.67</td>
<td>(12.216, 12.541)</td>
<td>0.066</td>
</tr>
<tr>
<td>0.21</td>
<td>(1.411, 2.128)</td>
<td>0.152</td>
<td>6.71</td>
<td>(12.299, 12.580)</td>
<td>0.064</td>
</tr>
<tr>
<td>0.48</td>
<td>(1.924, 2.587)</td>
<td>0.130</td>
<td>6.90</td>
<td>(12.658, 13.053)</td>
<td>0.079</td>
</tr>
<tr>
<td>0.81</td>
<td>(2.532, 4.039)</td>
<td>0.209</td>
<td>7.07</td>
<td>(13.019, 13.385)</td>
<td>0.070</td>
</tr>
<tr>
<td>1.21</td>
<td>(3.260, 4.329)</td>
<td>0.208</td>
<td>7.82</td>
<td>(14.387, 14.776)</td>
<td>0.080</td>
</tr>
<tr>
<td>1.69</td>
<td>(4.090, 5.401)</td>
<td>0.245</td>
<td>8.94</td>
<td>(16.322, 16.596)</td>
<td>0.066</td>
</tr>
<tr>
<td>2.25</td>
<td>(5.056, 6.130)</td>
<td>0.208</td>
<td>10.56</td>
<td>(18.912, 19.203)</td>
<td>0.059</td>
</tr>
<tr>
<td>2.90</td>
<td>(6.151, 7.178)</td>
<td>0.210</td>
<td>12.88</td>
<td>(22.540, 22.785)</td>
<td>0.047</td>
</tr>
<tr>
<td>3.64</td>
<td>(7.360, 8.833)</td>
<td>0.274</td>
<td>16.26</td>
<td>(27.601, 27.862)</td>
<td>0.047</td>
</tr>
<tr>
<td>4.48</td>
<td>(8.728, 10.087)</td>
<td>0.254</td>
<td>21.35</td>
<td>(35.217, 35.402)</td>
<td>0.042</td>
</tr>
<tr>
<td>5.42</td>
<td>(10.216, 10.994)</td>
<td>0.181</td>
<td>29.39</td>
<td>(46.961, 47.145)</td>
<td>0.037</td>
</tr>
<tr>
<td>6.48</td>
<td>(11.880, 12.907)</td>
<td>0.193</td>
<td>42.93</td>
<td>(66.497, 66.682)</td>
<td>0.034</td>
</tr>
<tr>
<td>6.60</td>
<td>(12.074, 12.314)</td>
<td>0.060</td>
<td>68.08</td>
<td>(102.316, 102.471)</td>
<td>0.032</td>
</tr>
<tr>
<td>6.63</td>
<td>(12.142, 12.385)</td>
<td>0.056</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 5.3. Loss vs. \( \nu \);
Method 1, \( f(x) \) as at (30)

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>loss</th>
<th>( \nu )</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.15</td>
<td>6.67</td>
<td>11.27</td>
</tr>
<tr>
<td>0.21</td>
<td>1.59</td>
<td>6.71</td>
<td>11.32</td>
</tr>
<tr>
<td>0.48</td>
<td>2.08</td>
<td>6.90</td>
<td>11.58</td>
</tr>
<tr>
<td>0.81</td>
<td>2.68</td>
<td>7.07</td>
<td>11.82</td>
</tr>
<tr>
<td>1.21</td>
<td>3.39</td>
<td>7.82</td>
<td>12.95</td>
</tr>
<tr>
<td>1.69</td>
<td>4.17</td>
<td>8.94</td>
<td>14.61</td>
</tr>
<tr>
<td>2.25</td>
<td>5.11</td>
<td>10.56</td>
<td>16.96</td>
</tr>
<tr>
<td>2.90</td>
<td>6.13</td>
<td>12.88</td>
<td>20.25</td>
</tr>
<tr>
<td>3.64</td>
<td>7.30</td>
<td>16.26</td>
<td>24.95</td>
</tr>
<tr>
<td>4.48</td>
<td>8.62</td>
<td>21.35</td>
<td>31.95</td>
</tr>
<tr>
<td>5.42</td>
<td>9.98</td>
<td>29.39</td>
<td>42.91</td>
</tr>
<tr>
<td>6.48</td>
<td>11.53</td>
<td>42.93</td>
<td>61.30</td>
</tr>
<tr>
<td>6.60</td>
<td>11.21</td>
<td>68.08</td>
<td>95.33</td>
</tr>
<tr>
<td>6.63</td>
<td>11.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.4. Interval of the loss (95%) and st. dev. using \( f(x) \) as at (30)

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>interval(95%)</th>
<th>Std. dev.</th>
<th>( \nu )</th>
<th>interval(95%)</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(1.000, 2.413)</td>
<td>0.231</td>
<td>6.67</td>
<td>(11.187, 11.363)</td>
<td>0.037</td>
</tr>
<tr>
<td>0.21</td>
<td>(1.359, 2.770)</td>
<td>0.292</td>
<td>6.71</td>
<td>(11.250, 11.379)</td>
<td>0.028</td>
</tr>
<tr>
<td>0.48</td>
<td>(1.825, 3.248)</td>
<td>0.256</td>
<td>6.90</td>
<td>(11.507, 11.669)</td>
<td>0.032</td>
</tr>
<tr>
<td>0.81</td>
<td>(2.395, 4.620)</td>
<td>0.326</td>
<td>7.07</td>
<td>(11.752, 11.906)</td>
<td>0.030</td>
</tr>
<tr>
<td>1.21</td>
<td>(3.003, 4.776)</td>
<td>0.368</td>
<td>7.82</td>
<td>(12.862, 13.056)</td>
<td>0.038</td>
</tr>
<tr>
<td>1.69</td>
<td>(3.755, 5.647)</td>
<td>0.310</td>
<td>8.94</td>
<td>(14.543, 14.710)</td>
<td>0.032</td>
</tr>
<tr>
<td>2.25</td>
<td>(4.655, 6.446)</td>
<td>0.346</td>
<td>10.56</td>
<td>(16.879, 17.026)</td>
<td>0.030</td>
</tr>
<tr>
<td>2.90</td>
<td>(5.629, 7.123)</td>
<td>0.312</td>
<td>12.88</td>
<td>(20.201, 20.299)</td>
<td>0.022</td>
</tr>
<tr>
<td>3.64</td>
<td>(6.762, 9.238)</td>
<td>0.397</td>
<td>16.26</td>
<td>(24.913, 24.982)</td>
<td>0.014</td>
</tr>
<tr>
<td>4.48</td>
<td>(8.021, 10.692)</td>
<td>0.516</td>
<td>21.35</td>
<td>(31.934, 31.964)</td>
<td>0.007</td>
</tr>
<tr>
<td>5.42</td>
<td>(9.431, 11.249)</td>
<td>0.370</td>
<td>29.39</td>
<td>(42.906, 42.915)</td>
<td>0.002</td>
</tr>
<tr>
<td>6.48</td>
<td>(10.956, 12.696)</td>
<td>0.429</td>
<td>42.93</td>
<td>(61.276, 61.320)</td>
<td>0.010</td>
</tr>
<tr>
<td>6.60</td>
<td>(11.116, 11.299)</td>
<td>0.039</td>
<td>68.08</td>
<td>(95.286, 95.367)</td>
<td>0.020</td>
</tr>
<tr>
<td>6.63</td>
<td>(11.143, 11.350)</td>
<td>0.041</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.1. Loss vs. $\nu$; Method 2, $f_0(x)$

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>loss</th>
<th>$\nu$</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.03</td>
<td>6.67</td>
<td>12.20</td>
</tr>
<tr>
<td>0.21</td>
<td>1.45</td>
<td>6.71</td>
<td>12.27</td>
</tr>
<tr>
<td>0.48</td>
<td>1.97</td>
<td>6.90</td>
<td>12.62</td>
</tr>
<tr>
<td>0.81</td>
<td>2.58</td>
<td>7.07</td>
<td>12.93</td>
</tr>
<tr>
<td>1.21</td>
<td>3.31</td>
<td>7.82</td>
<td>14.22</td>
</tr>
<tr>
<td>1.69</td>
<td>4.15</td>
<td>8.94</td>
<td>16.06</td>
</tr>
<tr>
<td>2.25</td>
<td>5.12</td>
<td>10.56</td>
<td>18.63</td>
</tr>
<tr>
<td>2.90</td>
<td>6.22</td>
<td>12.88</td>
<td>22.22</td>
</tr>
<tr>
<td>3.64</td>
<td>7.44</td>
<td>16.26</td>
<td>27.35</td>
</tr>
<tr>
<td>4.48</td>
<td>8.81</td>
<td>21.35</td>
<td>34.96</td>
</tr>
<tr>
<td>5.42</td>
<td>10.30</td>
<td>29.39</td>
<td>46.83</td>
</tr>
<tr>
<td>6.48</td>
<td>11.98</td>
<td>42.93</td>
<td>66.60</td>
</tr>
<tr>
<td>6.60</td>
<td>12.07</td>
<td>68.08</td>
<td>102.90</td>
</tr>
<tr>
<td>6.63</td>
<td>12.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6 Method 2

As discussed in the section one, we solve the following equation to get the design points

$$\xi_0(x) = \frac{i - 1}{n - 1}$$  \hspace{1cm} (31)

where $\xi_0(x)$ is defined in (28) and (29). This approximates the optimal $\xi_0(x)$ by a, necessarily symmetric, discrete measure.

Solving the equation (31) for $i = 1, 2, 3, ......., n$ where $n = 25$ by S-PLUS function “UNIROOT”, we get 25 design points. To find the loss function defined in (14) we compute $f_0(x), \tau_n$ and $\gamma_n$ for those design points. Table 6.1 gives numerical results of the loss function for different values of $\nu$.

For the function (30) Table 6.2 gives numerical results of the loss function for different values of $\nu$. 


Table 6.2. Loss vs. $\nu$; Method 2, $f(x)$ as at (30)

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>loss</th>
<th>$\nu$</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.04</td>
<td>6.67</td>
<td>11.53</td>
</tr>
<tr>
<td>0.21</td>
<td>1.45</td>
<td>6.71</td>
<td>11.57</td>
</tr>
<tr>
<td>0.48</td>
<td>1.97</td>
<td>6.90</td>
<td>11.83</td>
</tr>
<tr>
<td>0.81</td>
<td>2.58</td>
<td>7.07</td>
<td>12.08</td>
</tr>
<tr>
<td>1.21</td>
<td>3.30</td>
<td>7.82</td>
<td>13.23</td>
</tr>
<tr>
<td>1.69</td>
<td>4.12</td>
<td>8.94</td>
<td>14.93</td>
</tr>
<tr>
<td>2.25</td>
<td>5.07</td>
<td>10.56</td>
<td>17.33</td>
</tr>
<tr>
<td>2.90</td>
<td>6.14</td>
<td>12.88</td>
<td>20.69</td>
</tr>
<tr>
<td>3.64</td>
<td>7.31</td>
<td>16.26</td>
<td>25.49</td>
</tr>
<tr>
<td>4.48</td>
<td>8.62</td>
<td>21.35</td>
<td>32.66</td>
</tr>
<tr>
<td>5.42</td>
<td>10.06</td>
<td>29.39</td>
<td>43.87</td>
</tr>
<tr>
<td>6.48</td>
<td>11.66</td>
<td>42.93</td>
<td>62.70</td>
</tr>
<tr>
<td>6.60</td>
<td>11.46</td>
<td>68.08</td>
<td>97.53</td>
</tr>
<tr>
<td>6.63</td>
<td>11.48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7 Method 3

In this method, we solve the following equations for $x$,

$$\frac{1}{n} \sum_{i=1}^{n} x_i^{2k} = E_{\xi_0}(X^{2k})$$

(32)
for as many integral values of $k$ as possible. Then the empirical design measure tends to the true optimal design. For $\gamma \in [\frac{1}{12}, \frac{3}{20}]$,

\[
E_{\xi_0}(X^{2k}) = \int_{s} x^{2k} d\xi_0(x)
\]

\[
= \int_{\frac{1}{\gamma}}^{\frac{1}{\gamma}} x^{2k} m_0(x) \, dx
\]

\[
= \int_{\frac{1}{\gamma}}^{\frac{1}{\gamma}} x^{2k} \left[ 1 + \frac{5}{4}(12\gamma - 1)(12x^2 - 1) \right] \, dx
\]

\[
= \frac{1}{(2k+1)2^{2k}} \left[ \frac{60k\gamma - 3k + 3}{2k + 3} \right].
\]

Again for $\gamma \in [\frac{3}{20}, \frac{1}{4}]$, then

\[
E_{\xi_0}(X^{2k}) = \int_{s} x^{2k} d\xi_0(x)
\]

\[
= \int_{\frac{1}{\gamma}}^{\frac{1}{\gamma}} x^{2k} \left[ \frac{3}{1-c^2(1+2c)(4x^2-c^2)^+} \right] \, dx
\]

\[
= \frac{6}{(1-c)^2(1+2c)} \int_{\frac{1}{\gamma}}^{\frac{1}{\gamma}} x^{2k} (4x^2-c^2) \, dx
\]

\[
= \frac{6}{(1-c)^2(1+2c)} \left[ \frac{1}{(2k+3)2^{2k}} (1-c^{2k+3}) \right.
\]

\[
- \left. \frac{c^2}{(2k+1)2^{2k}} (1-c^{2k+1}) \right].
\]

From (32) we get, for $\gamma \in [\frac{1}{12}, \frac{3}{20}]$

\[
\frac{1}{n} \sum_{i=1}^{n} x_i^{2k} = \frac{1}{(2k+1)2^{2k}} \left[ \frac{60k\gamma - 3k + 3}{2k + 3} \right],
\]

(33)
and, for $\gamma \in \left[\frac{3}{20}, \frac{1}{4}\right]$ 

$$
\frac{1}{n} \sum_{i=1}^{n} x_i^{2k} = \frac{6}{(1 - c)^2(1 + 2c)} \left[ \frac{1}{(2k + 3)2^{2k}} (1 - c^{2k+3}) - \frac{c^2}{(2k + 1)2^{2k}} (1 - c^{2k+1}) \right].
$$

Above equations (33) and (34) are nonlinear system of equation for different integral values of $i$ and $k$. To solve these nonlinear systems, we use MATLAB function “FMINCON” which finds the constrained minimum of a scalar function of several variables starting at an initial value. This is generally referred to as “constrained nonlinear optimization”. In this method, we assume, design points $n = 25$ for $\nu \in [0, 6.48]$ and $n = 25$ for $\nu \in [6.48, \infty)$. To find the loss function defined in (14) we compute $f_0(x)$, $\tau_n$ and $\gamma_n$ for those design points. For a symmetric design, we solve the equations (33) and (34) for $x_1, x_2, x_3, \ldots, x_{\lfloor n/2 \rfloor}$ and then our designs are $0, \pm x_1, \pm x_2, \pm x_3, \cdots, \pm x_{\lfloor n/2 \rfloor}$. Now if $n = 25$, we solve the equations (33) and (34) for $k = 1, 2, 3, \cdots, 12$, and the odd moments are zero by symmetry. Table 7.1 gives numerical results of the loss function for different values of $\nu$.

For the function (30) Table 7.2 gives numerical results of the loss function for different values of $\nu$.

Table 7.3 represents the design points for a straight line regression using each of the methods 1-3.
Table 7.1. Loss vs. $\nu$;
Method 3, $f_0(x)$

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>loss</th>
<th>$\nu$</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>6.67</td>
<td>12.19</td>
</tr>
<tr>
<td>0.21</td>
<td>1.42</td>
<td>6.71</td>
<td>12.25</td>
</tr>
<tr>
<td>0.48</td>
<td>1.93</td>
<td>6.90</td>
<td>12.54</td>
</tr>
<tr>
<td>0.81</td>
<td>2.54</td>
<td>7.07</td>
<td>12.81</td>
</tr>
<tr>
<td>1.21</td>
<td>3.26</td>
<td>7.82</td>
<td>13.96</td>
</tr>
<tr>
<td>1.69</td>
<td>4.10</td>
<td>8.94</td>
<td>15.67</td>
</tr>
<tr>
<td>2.25</td>
<td>5.06</td>
<td>10.56</td>
<td>18.13</td>
</tr>
<tr>
<td>2.90</td>
<td>6.15</td>
<td>12.88</td>
<td>21.60</td>
</tr>
<tr>
<td>3.64</td>
<td>7.37</td>
<td>16.26</td>
<td>26.59</td>
</tr>
<tr>
<td>4.48</td>
<td>8.73</td>
<td>21.35</td>
<td>34.04</td>
</tr>
<tr>
<td>5.42</td>
<td>10.23</td>
<td>29.39</td>
<td>45.65</td>
</tr>
<tr>
<td>6.48</td>
<td>11.89</td>
<td>42.93</td>
<td>64.97</td>
</tr>
<tr>
<td>6.60</td>
<td>12.08</td>
<td>68.08</td>
<td>100.41</td>
</tr>
<tr>
<td>6.63</td>
<td>12.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2. Loss vs. $\nu$; Method 3, using $f(x)$ as at (30)

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>loss</th>
<th>$\nu$</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>6.67</td>
<td>11.52</td>
</tr>
<tr>
<td>0.21</td>
<td>1.41</td>
<td>6.71</td>
<td>11.73</td>
</tr>
<tr>
<td>0.48</td>
<td>1.92</td>
<td>6.90</td>
<td>12.01</td>
</tr>
<tr>
<td>0.81</td>
<td>2.50</td>
<td>7.07</td>
<td>12.29</td>
</tr>
<tr>
<td>1.21</td>
<td>3.19</td>
<td>7.82</td>
<td>13.38</td>
</tr>
<tr>
<td>1.69</td>
<td>3.99</td>
<td>8.94</td>
<td>15.02</td>
</tr>
<tr>
<td>2.25</td>
<td>4.89</td>
<td>10.56</td>
<td>17.38</td>
</tr>
<tr>
<td>2.90</td>
<td>5.91</td>
<td>12.88</td>
<td>20.73</td>
</tr>
<tr>
<td>3.64</td>
<td>6.94</td>
<td>16.26</td>
<td>25.61</td>
</tr>
<tr>
<td>4.48</td>
<td>8.32</td>
<td>21.35</td>
<td>32.88</td>
</tr>
<tr>
<td>5.42</td>
<td>9.71</td>
<td>29.39</td>
<td>44.63</td>
</tr>
<tr>
<td>6.48</td>
<td>11.27</td>
<td>42.93</td>
<td>63.35</td>
</tr>
<tr>
<td>6.60</td>
<td>10.97</td>
<td>68.08</td>
<td>98.37</td>
</tr>
<tr>
<td>6.63</td>
<td>11.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method 1</td>
<td>Method 2</td>
<td>Method 3</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>$\nu = 5.41$</td>
<td>$\nu = 68.08$</td>
<td>$\nu = 5.41$</td>
<td>$\nu = 68.08$</td>
</tr>
<tr>
<td>$\pm 0.493$</td>
<td>$\pm 0.499$</td>
<td>$\pm 0.485$</td>
<td>$\pm 0.495$</td>
</tr>
<tr>
<td>$\pm 0.453$</td>
<td>$\pm 0.495$</td>
<td>$\pm 0.468$</td>
<td>$\pm 0.489$</td>
</tr>
<tr>
<td>$\pm 0.426$</td>
<td>$\pm 0.482$</td>
<td>$\pm 0.451$</td>
<td>$\pm 0.482$</td>
</tr>
<tr>
<td>$\pm 0.421$</td>
<td>$\pm 0.479$</td>
<td>$\pm 0.432$</td>
<td>$\pm 0.476$</td>
</tr>
<tr>
<td>$\pm 0.399$</td>
<td>$\pm 0.486$</td>
<td>$\pm 0.412$</td>
<td>$\pm 0.469$</td>
</tr>
<tr>
<td>$\pm 0.361$</td>
<td>$\pm 0.476$</td>
<td>$\pm 0.389$</td>
<td>$\pm 0.461$</td>
</tr>
<tr>
<td>$\pm 0.361$</td>
<td>$\pm 0.451$</td>
<td>$\pm 0.364$</td>
<td>$\pm 0.453$</td>
</tr>
<tr>
<td>$\pm 0.345$</td>
<td>$\pm 0.449$</td>
<td>$\pm 0.335$</td>
<td>$\pm 0.444$</td>
</tr>
<tr>
<td>$\pm 0.275$</td>
<td>$\pm 0.454$</td>
<td>$\pm 0.300$</td>
<td>$\pm 0.433$</td>
</tr>
<tr>
<td>$\pm 0.194$</td>
<td>$\pm 0.462$</td>
<td>$\pm 0.255$</td>
<td>$\pm 0.421$</td>
</tr>
<tr>
<td>$\pm 0.168$</td>
<td>$\pm 0.407$</td>
<td>$\pm 0.189$</td>
<td>$\pm 0.404$</td>
</tr>
<tr>
<td>$\pm 0.055$</td>
<td>$\pm 0.421$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Figures 2-4 represent examples of designs for straight line regression using each of the methods 1-3.

Figure 2. Method 1; \( \nu = 5.42 \) and \( \nu = 68.08 \)

Figure 3. Method 2; \( \nu = 5.42 \) and \( \nu = 68.08 \)

Figure 4. Method 3; \( \nu = 5.42 \) and \( \nu = 68.08 \)
8 Conclusion

In this project we have implemented the design measure $\xi_0(x)$ for straight line regression by using a discrete approximation to the minimax design measure $\xi_0$. We have approached this implementation using three methods and we have made comparisons among these three methods using the MSE (Mean Square Error) as our overall measure of loss. In comparing these three methods, we observed that the maximum loss values are close to one when $\nu = 0$. For the method one, if we look at the interval containing of the loss values, then it is not wide that is, the variation of loss is not high. For all the methods, it is observed that the loss is monotonic as a function of $\nu$ for the least favourable $f_0(x)$ which is defined in (23) and not monotonic for the other $f(x)$ which is defined in (30).

From a computational point of view, Method 1 is simple. Method 2 is less complicated compared to the method 3. Method 3 is computationally harder than the other two methods because we apply the optimization technique to solve nonlinear system of equations in 12 unknowns.

9 References


