# Minimax Wavelet Designs for Curve Estimation

*Alwell J. Oyet*
*Memorial University of Newfoundland*
*Dept. of Mathematics and Statistics*
*St. John's, NF CANADA*
*aoyet@math.mun.ca*

## 1. INTRODUCTION

In this paper, we outline, the results of our investigation into using wavelets to construct A- and D-optimal integer-valued designs for estimating a nonparametric response. The approximately linear wavelet regression model

$$Y_{ij} \;=\; \mathbf{q}_m^T(x_i)\boldsymbol{\beta} + f(x_i) + \varepsilon_{ij}, i = 1, \cdots, N,\ j = 1, \cdots, n_i,\ \sum_{i=1}^{N} n_i = n; \tag{1}$$

$x_i \in [A_{i-1}, A_i];\ \ \bigcup_{i=1}^{N}[A_{i-1}, A_i] = [0,1]$ arises from representing the response function in a nonparametric regression model $\eta(x)$ by a finite *mth* order wavelet expansion. In (1), $\mathbf{q}_m(x)$ consists of $2^{m+1} \times 1$ dilated and translated versions $\psi^{-l,k}(x) = 2^{l/2}\psi(2^l x - k), (l,\ k$ integers) of a primary wavelet $\psi(x)$ and a scaling function $\phi(x)$. The term $f(x)$ which represents components of the wavelet system not used in the approximation will account for the uncertainty in the true structure of the response function. This term automatically introduces bias in the estimates of the response $\eta(x)$. In order to control the magnitude of the bias, we impose a bound $\frac{1}{N}\sum_{i=1}^{N} f^2(x_i) \le \tau^2$, on $f(x)$ for known constant $\tau$. In §2, we will use the discrete analogue of the identifiability condition on $\boldsymbol{\beta}$ $\frac{1}{N}\sum_{i=1}^{N}\mathbf{q}_m(x_i)f(x_i) = \mathbf{0}$ to obtain an expression for the contamination term $f(x)$. The error terms $\varepsilon_i$ are assumed to be uncorrelated with zero mean and constant variance $\sigma^2$. We assume that

$$\hat{\eta}(x) = \mathbf{q}_m^T(x)\hat{\boldsymbol{\beta}}_M = \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i} Y_{ij}\int_{A_{i-1}}^{A_i} K_m(x,s)ds; \tag{2}$$

where $\hat{\boldsymbol{\beta}}_M = \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i} Y_{ij}\mathbf{B}^{-1}\mathbf{z}_i$; $\mathbf{B} = \mathbf{Z}^T\mathbf{P}\mathbf{Q} = \sum_{i=1}^{N} p_i\mathbf{z}_i\mathbf{q}_m^T(x_i)$, $\mathbf{z}_i = \int_{A_{i-1}}^{A_i}\mathbf{q}_m(s)ds$ and $K_m(x,s) = \mathbf{q}_m^T(x)\mathbf{B}^{-1}\mathbf{q}_m(s)$. Here, we refer to $p_i = \frac{n_i}{n}$ as an integer-valued design (a probability distribution) on a finite but arbitrarily dense design space $\mathcal{S}$ constructed from the interval $[0,1]$. It can be verified that $\mathbf{d} = bias(\hat{\boldsymbol{\beta}}_M) = \mathbf{B}^{-1}\mathbf{b}$, where $\mathbf{b} = \mathbf{Z}^T\mathbf{P}\mathbf{f} = \sum_{i=1}^{N} p_i\mathbf{z}_i f(x_i)$ and $\mathbf{f} = (f(x_1), \cdots, f(x_N))^T$ with covariance matrix $\mathbf{C} = \frac{\sigma^2}{n}\mathbf{H}^{-1}$, where $\mathbf{H} = \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{A} = \mathbf{Z}^T\mathbf{P}\mathbf{Z} = \sum_{i=1}^{N} p_i\mathbf{z}_i\mathbf{z}_i^T$. A second method which the experimenter may consider for estimating $\boldsymbol{\beta}$ is by weighted least squares $\hat{\boldsymbol{\beta}}_{WLS} = \left[\frac{1}{n}\sum_{i=1}^{N} n_i w_i\mathbf{q}_m(x_i)\mathbf{q}_m^T(x_i)\right]^{-1}\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n_i} w_i\mathbf{q}_m(x_i)Y_{ij}$. Define $m_i = p_i w_i$, $\mathbf{B}_* = \mathbf{Q}^T\mathbf{M}\mathbf{Q}$; $\mathbf{b}_* = \mathbf{Q}^T\mathbf{M}\mathbf{f}$; $\mathbf{D}_* = \mathbf{Q}^T\mathbf{M}\mathbf{W}\mathbf{Q}$, where $\mathbf{M} = diag(m_1, \cdots, m_N)$, $\mathbf{W} = diag(w_1, \cdots, w_N)$ and $w_i = \frac{\int_0^1 \|\mathbf{q}(s)\|ds}{\|\mathbf{q}(x_i)\|}$. Then, $\mathbf{d}_* = bias(\hat{\boldsymbol{\beta}}_{WLS}) = \mathbf{B}_*^{-1}\mathbf{b}_*$ and $\mathbf{C}_* = Var(\hat{\boldsymbol{\beta}}_{WLS}) = \frac{\sigma^2}{n}\mathbf{H}_*^{-1}$, where $\mathbf{H}_* = \mathbf{B}_*\mathbf{D}_*^{-1}\mathbf{B}_*$.

## 2. MINIMAX INTEGER-VALUED DESIGNS

Let us assume that the $N \times r$ matrices $\mathbf{Q}$ and $\mathbf{Z}$ are of full rank. Define $\mathbf{Q} = \mathbf{U}_{N\times r}\boldsymbol{\Lambda}_{r\times r}\mathbf{V}_{r\times r}^T$, where $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = I_r$ and $\boldsymbol{\Lambda}$ is a diagonal matrix of singular values $\lambda_i(\mathbf{Q})$ $(i = 1, \cdots, r)$ of $\mathbf{Q}$. Similarly, we define $\mathbf{Z} = \mathbf{U}_{N\times r}^*\boldsymbol{\Lambda}_{r\times r}^*\mathbf{V}_{r\times r}^{*T}$. From the identifiability condition we observe that $\mathbf{V}\boldsymbol{\Lambda}\mathbf{U}^T\mathbf{f} = \mathbf{0}$. This suggests that $\mathbf{f}$ belongs to the space of orthogonal complements of $\mathbf{U}$ denoted by $[col(\mathbf{U})]^{\perp}$. That is, $\mathbf{f} = \tau\sqrt{N}\tilde{\mathbf{U}}\mathbf{c}$, for some $\mathbf{c}$ such that $\|\mathbf{c}\| \le 1$ and $\tilde{\mathbf{U}}_{N\times N-r}$ a matrix whose columns are elements of $[col(\mathbf{U})]^{\perp}$. It can be verified that if $\mathbf{P}_1 = \mathbf{U}^{*T}\mathbf{P}\mathbf{U}$; $\mathbf{P}_j^* = \mathbf{U}^{*T}\mathbf{P}^j\mathbf{U}^*$; $\tilde{\mathbf{P}}_1 = \mathbf{P}_1^T\mathbf{P}_1^{*-1}\mathbf{P}_1$; $\mathbf{M}_j = \mathbf{U}^T\mathbf{M}^j\mathbf{U}$; $\mathbf{M}^* = \mathbf{U}^T\mathbf{M}\mathbf{W}\mathbf{U}$ and $\tilde{\mathbf{M}}_1 = \mathbf{M}_1\mathbf{M}^{*-1}\mathbf{M}_1$, then Theorem 2.1 hold.

**Theorem 2..1** *Let $\lambda_{max}(\mathbf{G})$ be the maximum eigenvalue of $\mathbf{G} = \mathbf{P}_1^{*-1}(\mathbf{P}_2^* - \mathbf{P}_1\mathbf{P}_1^T)$ with corresponding eigenvector $\mathbf{c}$ and $\lambda_i(\tilde{\mathbf{P}}_1)$ $(i = 1, \cdots, r)$ be the eigenvalues of $\tilde{\mathbf{P}}_1$. Let $\lambda_{max}(\mathbf{G}_1)$ be the maximum eigenvalue of $\mathbf{G}_1 = \Lambda^{-1}(\mathbf{P}_1^{-1}\mathbf{P}_2^*(\mathbf{P}_1^T)^{-1} - \mathbf{I}_{r\times r})\Lambda^{-1}$ and $l_i$ be the diagonal elements of $\tilde{\mathbf{P}}_1^{-1}$. Then, for fixed $\nu = \frac{\sigma^2}{n\tau^2}$*

$$\max_f |\mathbf{MSE}(\hat{\boldsymbol{\beta}}_M)| = \left(\frac{\sigma^2}{n}\right)^r \frac{(1 + \nu^{-1}N\lambda_{max}(\mathbf{G}))}{\Pi_{i=1}^r \lambda_i^2(\mathbf{Q}) \cdot \lambda_i(\tilde{\mathbf{P}}_1)}.$$

$$\max_f \; tr\left\{\mathbf{MSE}(\hat{\boldsymbol{\beta}}_M)\right\} = \tau^2 \left[N\lambda_{max}(\mathbf{G}_1) + \nu \sum_{i=1}^r \lambda_i^{-2}(\mathbf{Q}) \cdot l_i\right].$$

*The expression for $\max_f |\mathbf{MSE}(\hat{\boldsymbol{\beta}}_{WLS})|$ is obtained if we replace $\tilde{\mathbf{P}}_1$ by $\tilde{\mathbf{M}}_1$ and $\lambda_{max}(\mathbf{G})$ becomes the maximum eigenvalue of $\mathbf{G} = \mathbf{M}^{*-1}(\mathbf{M}_2 - \mathbf{M}_1^2)$. Replace $\tilde{\mathbf{P}}_1$ by $\tilde{\mathbf{M}}_1$; $\mathbf{P}_1$ by $\mathbf{M}_1$ and $\mathbf{P}_2^*$ by $\mathbf{M}_2$, to obtain the expression for $\max_f \; tr\left\{\mathbf{MSE}(\hat{\boldsymbol{\beta}}_{WLS})\right\}$.*

In Figure 1, we display a design for estimating the response in an experiment on the utilization of nitrite in bush beans described in Bates and Watts (1988). The design was constructed using the simulated annealing algorithm; see Haines (1987).

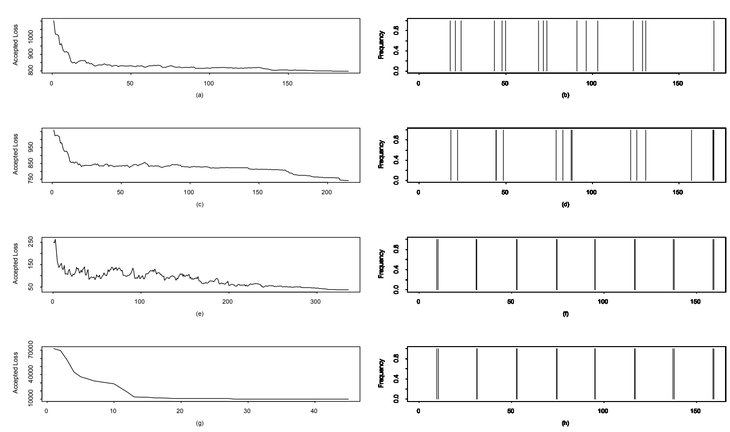

Figure 1: Exact integer-valued designs for nitrite utilization experiment. Daubechies wavelet ($N = 608$, $n = 16$, $m = 2$, $\nu = 0.05$): (a),(b) accepted loss and A-optimal design under Modified GM estimation (2); (c),(d) accepted loss and A-optimal design under WLS estimation; (e), (f) accepted loss and D-optimal design under Modified GM estimation (2); (g), (h) accepted loss and D-optimal design under WLS estimation.

## REFERENCES


Bates, D. M., and Watts, D. G. (1988), *Nonlinear Regression Analysis & its Applications*, New York: John Wiley and Sons Inc.

Haines, L. M. (1987), "The Application of the Annealing Algorithm to the Construction of Exact Optimal Designs for Linear-Regression Models," *Technometrics*, 29, 439-447.


## RESUME


En cet article, nous traçons les grandes lignes des résultats de notre recherche sur employer des wavelets pour construire les conceptions D-optimales de A et nombre-évalué pour estimer une réponse non paramétrique basée sur l'algorithme de recuit simulé; voir le Haines (1987). Le schéma 1 est une illustration d'une conception pour estimer la réponse dans une expérience sur l'utilisation du nitrite dans des haricots de buisson; voir les Bates et le Watts (1988).