University of Alberta
Department of Mathematical and Statistical Sciences
Technical Report S129

ROBUST QUANTILE REGRESSION DESIGNS

Linglong Kong and Douglas P. Wiens[1]

September 17, 2014

Abstract This technical report contains unpublished material, relevant to the article 'Model-Robust Designs for Quantile Regression'.

1 Proof of Theorem 1

The 'true' parameter $\boldsymbol{\theta}$ is defined by

$$\mathbf{0} = \int_{\chi} E_{Y|\boldsymbol{x}} \left[\psi_\tau \left(Y - \boldsymbol{f}' \left(\boldsymbol{x} \right) \boldsymbol{\theta} \right) \right] \boldsymbol{f} \left(\boldsymbol{x} \right) d\boldsymbol{x}. \tag{B.1}$$

The estimate is defined by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{t}} \sum_{i=1}^{n} \rho_\tau \left(Y_i - \boldsymbol{f}' \left(\boldsymbol{x}_i \right) \boldsymbol{t} \right), \tag{B.2}$$

where $\rho_\tau(\cdot)$ is the 'check' function $\rho_\tau(r) = r(\tau - I(r<0))$, with derivative $\psi_\tau(r) = \tau - I(r<0)$. Define the target parameter $\boldsymbol{\theta}$ to be the asymptotic solution to (B.2), so that

$$\sum_{i=1}^{n} \xi_{n,i} \psi_\tau \left(Y_i - \boldsymbol{f}' \left(\boldsymbol{x}_i \right) \boldsymbol{\theta} \right) \boldsymbol{f} \left(\boldsymbol{x}_i \right) \xrightarrow{pr} \mathbf{0}, \tag{B.3}$$

in agreement with (B.1). We require the following conditions.

(A1) The distribution function G_ε defined on $(-\infty, \infty)$ is twice continuously differentiable. The density g_ε is everywhere finite, positive and Lipschitz continuous.

(A2) $\max_{i=1,\ldots,n} \frac{1}{\sqrt{n}} ||\boldsymbol{f}(\boldsymbol{x}_i)|| \to 0$.

[1]Department of Mathematical and Statistical Sciences; University of Alberta, Edmonton, Alberta; Canada T6G 2G1. e-mail: `lkong@ualberta.ca, doug.wiens@ualberta.ca`

(A3) There exists a vector $\boldsymbol{\mu}$, and positive definite matrices $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$, such that, with $\delta_n^*(\boldsymbol{x}) = \delta_n(\boldsymbol{x})/\sigma(\boldsymbol{x})$,

$$\boldsymbol{\mu} = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\tau - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right) \boldsymbol{f}(\boldsymbol{x}_i),$$

$$\boldsymbol{\Sigma}_0 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i)) \left(1 - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right) \boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i),$$

$$\boldsymbol{\Sigma}_1 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{g_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))}{\sigma(\boldsymbol{x}_i)} \boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i).$$

Recall the definitions

$$\boldsymbol{\mu}_0 = \int_\chi \delta_0(\boldsymbol{x}) \frac{1}{\sigma(\boldsymbol{x})} \boldsymbol{f}(\boldsymbol{x}) \xi_\infty(d\boldsymbol{x}), \tag{B.4a}$$

$$\boldsymbol{P}_0 = \int_\chi \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x}) \xi_\infty(d\boldsymbol{x}), \tag{B.4b}$$

$$\boldsymbol{P}_1 = \int_\chi \boldsymbol{f}(\boldsymbol{x}) \frac{1}{\sigma(\boldsymbol{x})} \boldsymbol{f}'(\boldsymbol{x}) \xi_\infty(d\boldsymbol{x}). \tag{B.4c}$$

Assume that the support of ξ_∞ is large enough that $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ are positive definite. We have:

Theorem 1 *Under conditions (A1) – (A3) the quantile regression estimate $\hat{\boldsymbol{\theta}}_n$ of the parameter $\boldsymbol{\theta}$ defined by (B.3) is asymptotically normally distributed:*

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \xrightarrow{L} N\left(\boldsymbol{P}_1^{-1}\boldsymbol{\mu}_0, \frac{\tau(1-\tau)}{g_\varepsilon^2(0)} \boldsymbol{P}_1^{-1}\boldsymbol{P}_0\boldsymbol{P}_1^{-1}\right). \tag{B.5}$$

Proof Here we write an n-point design as $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$, with the $\boldsymbol{x}_i \in \chi$ not necessarily distinct. We first show that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{L} N(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}, \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}). \tag{B.6}$$

For this, define $Z_n(\boldsymbol{\gamma}) = \sum_{i=1}^n \left(\rho_\tau(u_i - f'(\boldsymbol{x}_i)\boldsymbol{\gamma}'\sqrt{n}) - \rho_\tau(u_i)\right)$, where $u_i = Y_i - \boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{\theta}$ and $\hat{\boldsymbol{\gamma}} = \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$. The function $Z_n(\boldsymbol{\gamma})$ is convex and is minimized at $\hat{\boldsymbol{\gamma}}$. The main idea of the proof follows Knight (1998). Using Knight's identity

$$\rho_\tau(u+v) - \rho_\tau(u) = -v\psi_\tau(u) + \int_0^v \left(I(u \le s) - I(u \le 0)\right) ds,$$

we may write $Z_n(\gamma) = Z_{1n}(\gamma) + Z_{2n}(\gamma)$, where

$$Z_{1n}(\boldsymbol{\gamma}) = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{f}'(\boldsymbol{x}_i)\gamma\psi_\tau(u_i),$$

$$Z_{2n}(\boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^{v_{ni}} \left(I(u_i \le s) - I(u_i \le 0)\right) ds \stackrel{def}{=} \sum_{i=1}^{n} Z_{2ni}(\boldsymbol{\gamma}),$$

and $v_{ni} = \boldsymbol{\gamma}'\boldsymbol{f}(\boldsymbol{x}_i)\sqrt{n}$. We note that

$$E[Z_{1n}(\boldsymbol{\gamma})] = -\boldsymbol{\gamma}'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{f}(\boldsymbol{x}_i)E[\psi_\tau(u_i)] = -\boldsymbol{\gamma}'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\tau - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right)\boldsymbol{f}(\boldsymbol{x}_i)$$

and that

$$\begin{aligned}\text{VAR}[Z_{1n}(\boldsymbol{\gamma})] &= \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i)\text{VAR}[\psi_\tau(u_i)]\boldsymbol{\gamma} \\ &= \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\left(1 - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right)\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{\gamma}.\end{aligned}$$

It follows from the Lindeberg-Feller Central Limit Theorem, using Condition (A3), that $Z_{1n}(\boldsymbol{\gamma}) \xrightarrow{L} -\boldsymbol{\gamma}'\boldsymbol{w}$ where $\boldsymbol{w} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_0)$. Now centre $Z_{2n}(\boldsymbol{\gamma})$:

$$Z_{2n}(\boldsymbol{\gamma}) = \sum E[Z_{2ni}(\boldsymbol{\gamma})] + \sum\left(Z_{2ni}(\boldsymbol{\gamma}) - E[Z_{2ni}(\boldsymbol{\gamma})]\right).$$

We have

$$\begin{aligned}\sum E[Z_{2ni}(\boldsymbol{\gamma})] &= \sum\int_0^{v_{ni}}\left(G_\varepsilon\left(-\delta_n^*(\boldsymbol{x}_i) + \frac{s}{\sigma(\boldsymbol{x}_i)}\right) - G_\varepsilon\left(-\delta_n^*(\boldsymbol{x}_i)\right)\right)ds \\ &= \frac{1}{n}\sum\int_0^{\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{\gamma}} g_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\frac{t}{\sigma(\boldsymbol{x}_i)}dt + o(1) \\ &= \frac{1}{2n}\sum\frac{g_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))}{\sigma(\boldsymbol{x}_i)}\boldsymbol{\gamma}'\boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{\gamma} + o(1) \\ &\to \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_1\boldsymbol{\gamma}.\end{aligned}$$

As well, we have the bound

$$\begin{aligned}\text{VAR}[Z_{2n}(\boldsymbol{\gamma})] &\leq \sum E\left[\int_0^{v_{ni}}\left(I(u_i \leq s) - I(u_i \leq 0)\right)ds\right]^2 \\ &\leq \sum E\left[\int_0^{v_{ni}}ds\int_0^{v_{ni}}\left(I(u_i \leq s) - I(u_i \leq 0)\right)ds\right] \\ &= \sum E\left[\frac{1}{\sqrt{n}}\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{\gamma}\int_0^{v_{ni}}\left(I(u_i \leq s) - I(u_i \leq 0)\right)ds\right] \\ &\leq \frac{1}{\sqrt{n}}\max|\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{\gamma}|E[Z_{2n}(\boldsymbol{\gamma})].\end{aligned}$$

Condition (A2) implies that $\text{VAR}[Z_{2n}(\boldsymbol{\gamma})] \to 0$. As a consequence, $\sum\left(Z_{2ni}(\boldsymbol{\gamma}) - E[Z_{2ni}(\boldsymbol{\gamma})]\right) \xrightarrow{pr} 0$ and $Z_{2n}(\boldsymbol{\gamma}) \xrightarrow{pr} \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_1\boldsymbol{\gamma}$. Combining these observations, we have

$$Z_n(\boldsymbol{\gamma}) \xrightarrow{L} Z_0(\boldsymbol{\gamma}) = -\boldsymbol{\gamma}'\boldsymbol{w} + \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_1\boldsymbol{\gamma}.$$

The convexity of the limiting objective function $Z_0(\boldsymbol{\gamma})$ ensures the uniqueness of the minimizer, which is $\boldsymbol{\gamma}_0 = \boldsymbol{\Sigma}_1^{-1}\boldsymbol{w}$. Therefore, we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \hat{\boldsymbol{\gamma}} = \arg\min Z_n(\boldsymbol{\gamma}) \xrightarrow{L} \boldsymbol{\gamma}_0 = \arg\min Z_0(\boldsymbol{\gamma}). \tag{B.7}$$

Similar arguments can be found in Pollard (1991) and Knight (1998). From (B.7) we immediately obtain (B.6).

To go from (B.6) to (B.5) requires passing from the limits in (A3) to (B.4). The expansion

$$\frac{1}{\sqrt{n}}\left(\tau - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right) = \frac{1}{n}\sqrt{n}\left(G_\varepsilon(0) - G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\right) = \frac{1}{n}\left(g_\varepsilon(0)\delta_0^*(\boldsymbol{x}_i) + o(1)\right)$$

yields $\boldsymbol{\mu} = g_\varepsilon(0)\boldsymbol{\mu}_0$. Here we require $\lim \frac{1}{n}\sum_{i=1}^n \boldsymbol{f}(\boldsymbol{x}_i)$ to be bounded; this is implied by the existence of $\boldsymbol{P}_0 = \lim \int_\chi \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x})\xi_n(d\boldsymbol{x})$:

$$\|\frac{1}{n}\sum_{i=1}^n \boldsymbol{f}(\boldsymbol{x}_i)\|^2 \le \frac{1}{n}\sum_{i=1}^n \|\boldsymbol{f}(\boldsymbol{x}_i)\|^2 = \frac{1}{n}\sum_{i=1}^n tr[\boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)] = tr[\frac{1}{n}\sum_{i=1}^n \boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)] \to tr\boldsymbol{P}_0.$$

Similarly, the expansion $G_\varepsilon(-\delta_n^*(\boldsymbol{x}_i)) = G_\varepsilon(0) - O(n^{-1/2}) = \tau - O(n^{-1/2})$ gives that

$$\boldsymbol{\Sigma}_0 = \lim\left\{\tau(1-\tau)\int_\chi \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x})\xi_n(d\boldsymbol{x}) + O(n^{-1/2})\right\} = \tau(1-\tau)\boldsymbol{P}_0.$$

Finally, the expansion $g_\varepsilon(-\delta_n^*(\boldsymbol{x}_i)) = g_\varepsilon(0) + o(n^{-1/2})$ gives

$$\boldsymbol{\Sigma}_1 = \lim \frac{1}{n}\sum_{i=1}^n g_\varepsilon(-\delta_n^*(\boldsymbol{x}_i))\boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)/\sigma(\boldsymbol{x}_i) = g_\varepsilon(0)\boldsymbol{P}_1. \qquad \square$$

2 Variance functions $\sigma_\xi^2(\boldsymbol{x})$ - additional examples

We consider classes $\Sigma_0 = \{\sigma_\xi(\cdot|r) | r \in (-\infty,\infty)\}$ of variance functions given by

$$\sigma_\xi(\boldsymbol{x}_i|r) = \begin{cases} c_r\xi_i^{r/2}, & \xi_i > 0, \\ 0, & \xi_i = 0, \end{cases} \text{ with } c_r = \left(\frac{\sum_{\xi_i>0}\xi_i^r}{N}\right)^{-1/2}, \tag{B.8a}$$

$$\sigma_\xi(\boldsymbol{x}|r) = \begin{cases} c_r m^{r/2}(\boldsymbol{x}), & m(\boldsymbol{x}) > 0, \\ 0, & m(\boldsymbol{x}) = 0, \end{cases} \text{ with } c_r = \left(\int_{m(\boldsymbol{x})>0} m^r(\boldsymbol{x})\,d\boldsymbol{x}\right)^{-1/2}, \tag{B.8b}$$

in discrete and continuous spaces respectively. When the experimenter seeks protection against a *fixed* alternative to homoscedasticity, i.e. fixed r, some cases of (B.8) may be treated in generality.

Under (B.8a) the maximized loss $\mathcal{L}_\nu(\xi|\sigma) = (1-\nu)\,tr\,(\boldsymbol{A}\boldsymbol{T}_0) + \nu ch_{\max}(\boldsymbol{A}\boldsymbol{T}_2)$ is

$$\mathcal{L}_\nu(\xi|r) = (1-\nu)\,c_r^2 tr\left(\boldsymbol{A}\boldsymbol{S}_1^{-1}(r)\,\boldsymbol{S}_0\boldsymbol{S}_1^{-1}(r)\right) + \nu ch_{\max}\left(\boldsymbol{A}\boldsymbol{S}_1^{-1}(r)\,\boldsymbol{S}_2(r)\,\boldsymbol{S}_1^{-1}(r)\right), \tag{B.9}$$

where

$$\boldsymbol{S}_0 = \sum_{\xi_i>0} \boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)\xi_i,$$

$$\boldsymbol{S}_k = \boldsymbol{S}_k(r) = \sum_{\xi_i>0} \boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}'(\boldsymbol{x}_i)\xi_i^{k\left(1-\frac{r}{2}\right)} \text{ for } k = 1, 2.$$

Note that $\boldsymbol{S}_0 = \boldsymbol{S}_1(0) = \boldsymbol{S}_2(1)$.

2.1 Discrete designs for variance functions (B.8) with r fixed

Example 2.1. If $r = 2$ then $\boldsymbol{S}_1 = \boldsymbol{S}_2 = \boldsymbol{A}_\xi$ and

$$\mathcal{L}_\nu(\xi|r=2) = (1-\nu)N\frac{\sum_{i=1}^N \xi_i \boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{A}_\xi^{-1}\boldsymbol{A}\boldsymbol{A}_\xi^{-1}\boldsymbol{f}(\boldsymbol{x}_i)}{\sum_{i=1}^N \xi_i^2} + \nu ch_{\max}\left(\boldsymbol{A}\boldsymbol{A}_\xi^{-1}\right). \tag{B.10}$$

Without some restriction on the class of designs so as to make it compact, there are sequences $\{\xi_\beta\}$ of designs for which $\mathcal{L}_\nu(\xi_\beta)$ tends to the minimum value of (B.10) as $\beta \to 0$, but ξ_0 has one-point support, so that $\boldsymbol{A}_{\xi_0}$ is singular. To see this, define $s_0 = \min_{1\le i\le N}\left\{\boldsymbol{f}'(\boldsymbol{x}_i)\boldsymbol{A}^{-1}\boldsymbol{f}(\boldsymbol{x}_i)\right\}$. Since $\boldsymbol{A}_\xi^{-1} \succeq (N\boldsymbol{A})^{-1}$ and $\sum_{i=1}^N \xi_i^2 \le 1$, we have that $\mathcal{L}_\nu(\xi|r=2) \ge ((1-\nu)s_0+\nu)/N \overset{def}{=} \mathcal{L}_{\min}$. If ξ_β places mass $1-\beta$ at an $\boldsymbol{x}_*$ for which s_0 is attained, and mass $\beta/(N-1)$ at every other point $\boldsymbol{x}_i$, then $\boldsymbol{A}_{\xi_\beta} = N\boldsymbol{A}$ and so $\mathcal{L}_\nu(\xi_\beta) = \mathcal{L}_{\min} + O(\beta)$ as $\beta \to 0$. This degeneracy can be avoided by, for instance, imposing a positive lower bound on the non-zero design weights.

2.2 Continuous designs for variance functions (B.8b) with r fixed

Example 2.1 continued. If $r = 2$ then $\boldsymbol{S}_1 = \boldsymbol{S}_2 = \boldsymbol{A}_m$ and

$$\mathcal{L}_\nu(\xi|r=2) = (1-\nu)\frac{\int_\chi \boldsymbol{f}'(\boldsymbol{x})\boldsymbol{A}_m^{-1}\boldsymbol{A}\boldsymbol{A}_m^{-1}\boldsymbol{f}(\boldsymbol{x})m(\boldsymbol{x})d\boldsymbol{x}}{\int_\chi m^2(\boldsymbol{x})d\boldsymbol{x}} + \nu ch_{\max}\left(\boldsymbol{A}\boldsymbol{A}_m^{-1}\right).$$

As in the discrete version of this example, a degenerate solution can be avoided at the cost of imposing superfluous restrictions on the designs.

Example 2.2 $r = 1$. The case $r = 1$ and $c_1 = 1$ results in

$$\mathcal{L}_\nu(\xi|r=1) = (1-\nu)\,tr\left(\boldsymbol{A}\boldsymbol{S}_1^{-1}(1)\,\boldsymbol{S}_0\boldsymbol{S}_1^{-1}(1)\right) + \nu ch_{\max}\left(\boldsymbol{A}\boldsymbol{S}_1^{-1}(1)\,\boldsymbol{S}_0\boldsymbol{S}_1^{-1}(1)\right).$$

The optimal design is uniform, with density $m_*(\boldsymbol{x}) \equiv 1/\text{VOL}(\chi)$. To prove this we note that it is sufficient to show that $\boldsymbol{S}_1^{-1}(1)\,\boldsymbol{S}_0\boldsymbol{S}_1^{-1}(1) \succeq \boldsymbol{A}^{-1}$. This is established by introducing $\boldsymbol{A}_m = \int_{m(\boldsymbol{x})>0} \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x})d\boldsymbol{x}$ and then using Proposition 1 to obtain $\boldsymbol{S}_1^{-1}(1)\,\boldsymbol{S}_0\boldsymbol{S}_1^{-1}(1) \succeq \boldsymbol{A}_m^{-1} \succeq \boldsymbol{A}^{-1} = \boldsymbol{A}_{m_*}^{-1}$.

3 Calculations for the construction of continuous minimax designs for quadratic regression and fixed variance functions

We consider symmetric designs and variance functions: $m(x) = m(-x)$ and $\sigma(x) = \sigma(-x)$. In terms of

$$\mu_i = \int_{-1}^{1} x^i m(x) dx, \kappa_i = \int_{-1}^{1} x^i \frac{m(x)}{\sigma(x)} dx, \omega_i = \int_{-1}^{1} x^i \left(\frac{m(x)}{\sigma(x)}\right)^2 dx$$

we have that

$$\boldsymbol{T}_{0,0} = \begin{pmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & 0 \\ \mu_2 & 0 & \mu_4 \end{pmatrix}, \boldsymbol{T}_{0,1} = \begin{pmatrix} \kappa_0 & 0 & \kappa_2 \\ 0 & \kappa_2 & 0 \\ \kappa_2 & 0 & \kappa_4 \end{pmatrix}, \boldsymbol{T}_{0,2} = \begin{pmatrix} \omega_0 & 0 & \omega_2 \\ 0 & \omega_2 & 0 \\ \omega_2 & 0 & \omega_4 \end{pmatrix},$$

and

$$\boldsymbol{T}_{0,1}^{-1} = \frac{1}{(\kappa_4\kappa_0 - \kappa_2^2)} \begin{pmatrix} \kappa_4 & 0 & -\kappa_2 \\ 0 & \kappa_2^{-1} & 0 \\ -\kappa_2 & 0 & \kappa_0 \end{pmatrix}, \ \mathbf{A} = 2 \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} \end{pmatrix}.$$

Define $\pi = \pi(m) = 2\left(\kappa_4\kappa_0 - \kappa_2^2\right)^{-2}$. Then

$$\begin{aligned}
\boldsymbol{T}_2 &= \boldsymbol{T}_{0,1}^{-1} \boldsymbol{T}_{0,2} \boldsymbol{T}_{0,1}^{-1} \\
&= \frac{\pi}{2} \begin{pmatrix} \kappa_4 & 0 & -\kappa_2 \\ 0 & \kappa_2^{-1} & 0 \\ -\kappa_2 & 0 & \kappa_0 \end{pmatrix} \begin{pmatrix} \omega_0 & 0 & \omega_2 \\ 0 & \omega_2 & 0 \\ \omega_2 & 0 & \omega_4 \end{pmatrix} \begin{pmatrix} \kappa_4 & 0 & -\kappa_2 \\ 0 & \kappa_2^{-1} & 0 \\ -\kappa_2 & 0 & \kappa_0 \end{pmatrix} \\
&= \frac{\pi}{2} \begin{pmatrix} \kappa_4\omega_0 - \kappa_2\omega_2 & 0 & \kappa_4\omega_2 - \kappa_2\omega_4 \\ 0 & \kappa_2^{-1}\omega_2 & 0 \\ -\kappa_2\omega_0 + \kappa_0\omega_2 & 0 & -\kappa_2\omega_2 + \kappa_0\omega_4 \end{pmatrix} \begin{pmatrix} \kappa_4 & 0 & -\kappa_2 \\ 0 & \kappa_2^{-1} & 0 \\ -\kappa_2 & 0 & \kappa_0 \end{pmatrix} \\
&= \frac{\pi}{2} \begin{pmatrix} \begin{array}{c} \kappa_4(\kappa_4\omega_0 - \kappa_2\omega_2) \\ -\kappa_2(\kappa_4\omega_2 - \kappa_2\omega_4) \end{array} & 0 & \begin{array}{c} -\kappa_2(\kappa_4\omega_0 - \kappa_2\omega_2) \\ +\kappa_0(\kappa_4\omega_2 - \kappa_2\omega_4) \end{array} \\ 0 & \frac{\omega_2}{\kappa_2^2} & 0 \\ \begin{array}{c} \kappa_4(-\kappa_2\omega_0 + \kappa_0\omega_2) \\ -\kappa_2(-\kappa_2\omega_2 + \kappa_0\omega_4) \end{array} & 0 & \begin{array}{c} -\kappa_2(-\kappa_2\omega_0 + \kappa_0\omega_2) \\ +\kappa_0(-\kappa_2\omega_2 + \kappa_0\omega_4) \end{array} \end{pmatrix} \\
&= \frac{\pi}{2} \begin{pmatrix} \kappa_4^2\omega_0 - 2\kappa_4\kappa_2\omega_2 + \kappa_2^2\omega_4 & 0 & \begin{array}{c} -\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 \\ +(\kappa_4\kappa_0 + \kappa_2^2)\omega_2 \end{array} \\ 0 & \frac{\omega_2}{\kappa_2^2} & 0 \\ \begin{array}{c} -\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 \\ +(\kappa_4\kappa_0 + \kappa_2^2)\omega_2 \end{array} & 0 & \kappa_2^2\omega_0 - 2\kappa_2\kappa_0\omega_2 + \kappa_0^2\omega_4 \end{pmatrix};
\end{aligned}$$

hence (replacing ω_i by μ_i in the above)

$$\boldsymbol{T}_0 = \boldsymbol{T}_{0,1}^{-1}\boldsymbol{T}_{0,0}\boldsymbol{T}_{0,1}^{-1} = \frac{\pi}{2}\begin{pmatrix} \kappa_4^2 - 2\kappa_4\kappa_2\mu_2 + \kappa_2^2\mu_4 & 0 & \begin{array}{c}-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\end{array} \\ 0 & \frac{\mu_2}{\kappa_2^2} & 0 \\ \begin{array}{c}-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\end{array} & 0 & \kappa_2^2 - 2\kappa_2\kappa_0\mu_2 + \kappa_0^2\mu_4 \end{pmatrix}.$$

Then

$$\begin{aligned} tr\left(\mathbf{A}\boldsymbol{T}_0\right) &= \pi tr \begin{pmatrix} 1 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1/3 & 0 & 1/5 \end{pmatrix}\begin{pmatrix} \kappa_4^2 - 2\kappa_4\kappa_2\mu_2 + \kappa_2^2\mu_4 & 0 & \begin{array}{c}-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\end{array} \\ 0 & \frac{\mu_2}{\kappa_2^2} & 0 \\ \begin{array}{c}-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\end{array} & 0 & \kappa_2^2 - 2\kappa_2\kappa_0\mu_2 + \kappa_0^2\mu_4 \end{pmatrix} \\ &= \pi\left\{\begin{array}{c} \left[\kappa_4^2 - 2\kappa_4\kappa_2\mu_2 + \kappa_2^2\mu_4\right] + \frac{1}{3}\left[-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\right] \\ +\frac{1}{3}\frac{\mu_2}{\kappa_2^2} \\ \frac{1}{3}\left[-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\right] + \frac{1}{5}\left[\kappa_2^2 - 2\kappa_2\kappa_0\mu_2 + \kappa_0^2\mu_4\right] \end{array}\right\} \\ &= \pi\left\{\begin{array}{c} \left[\kappa_4^2 - 2\kappa_4\kappa_2\mu_2 + \kappa_2^2\mu_4\right] + \frac{2}{3}\left[-\kappa_4\kappa_2 - \kappa_2\kappa_0\mu_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\mu_2\right] \\ +\frac{1}{3}\frac{\mu_2}{\kappa_2^2} + \frac{1}{5}\left[\kappa_2^2 - 2\kappa_2\kappa_0\mu_2 + \kappa_0^2\mu_4\right] \end{array}\right\} \\ &= \pi\left\{\begin{array}{c} \left[\kappa_4^2 - \frac{2}{3}\kappa_4\kappa_2 + \frac{1}{5}\kappa_2^2\right] + \left[\frac{1}{3\kappa_2^2} - 2\kappa_4\kappa_2 + \frac{2}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{5}\kappa_2\kappa_0\right]\mu_2 \\ + \left[\kappa_2^2 - \frac{2}{3}\kappa_2\kappa_0 + \frac{1}{5}\kappa_0^2\right]\mu_4 \end{array}\right\} \\ &\stackrel{def}{=} \rho_0\left(m\right). \end{aligned}$$

We have that

$$\mathbf{A}\boldsymbol{T}_2 = \pi\begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} \end{pmatrix}\begin{pmatrix} \kappa_4^2\omega_0 - 2\kappa_4\kappa_2\omega_2 + \kappa_2^2\omega_4 & 0 & \begin{array}{c}-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\end{array} \\ 0 & \frac{\omega_2}{\kappa_2^2} & 0 \\ \begin{array}{c}-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4\\ +\left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\end{array} & 0 & \kappa_2^2\omega_0 - 2\kappa_2\kappa_0\omega_2 + \kappa_0^2\omega_4 \end{pmatrix}$$

$$= \pi\begin{pmatrix} \begin{array}{c}\left[\kappa_4^2\omega_0 - 2\kappa_4\kappa_2\omega_2 + \kappa_2^2\omega_4\right]\\ +\frac{1}{3}\left[-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\right]\end{array} & 0 & \begin{array}{c}\left[-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\right]\\ +\frac{1}{3}\left[\kappa_2^2\omega_0 - 2\kappa_2\kappa_0\omega_2 + \kappa_0^2\omega_4\right]\end{array} \\ 0 & \frac{\omega_2}{3\kappa_2^2} & 0 \\ \begin{array}{c}\frac{1}{3}\left[\kappa_4^2\omega_0 - 2\kappa_4\kappa_2\omega_2 + \kappa_2^2\omega_4\right]\\ +\frac{1}{5}\left[-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\right]\end{array} & 0 & \begin{array}{c}\frac{1}{3}\left[-\kappa_4\kappa_2\omega_0 - \kappa_2\kappa_0\omega_4 + \left(\kappa_4\kappa_0 + \kappa_2^2\right)\omega_2\right]\\ +\frac{1}{5}\left[\kappa_2^2\omega_0 - 2\kappa_2\kappa_0\omega_2 + \kappa_0^2\omega_4\right]\end{array} \end{pmatrix}$$

$$= \pi\begin{pmatrix} \begin{array}{c}\left[\kappa_4^2 - \frac{1}{3}\kappa_4\kappa_2\right]\omega_0\\ +\left[\frac{1}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - 2\kappa_4\kappa_2\right]\omega_2\\ +\left[\kappa_2^2 - \frac{1}{3}\kappa_2\kappa_0\right]\omega_4\end{array} & 0 & \begin{array}{c}\left[\frac{1}{3}\kappa_2^2 - \kappa_4\kappa_2\right]\omega_0\\ +\left[\kappa_4\kappa_0 + \kappa_2^2 - \frac{2}{3}\kappa_2\kappa_0\right]\omega_2\\ +\left[\frac{1}{3}\kappa_0^2 - \kappa_2\kappa_0\right]\omega_4\end{array} \\ 0 & \frac{\omega_2}{3\kappa_2^2} & 0 \\ \begin{array}{c}\left[\frac{1}{3}\kappa_4^2 - \frac{1}{5}\kappa_4\kappa_2\right]\omega_0\\ +\left[\frac{1}{5}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{3}\kappa_4\kappa_2\right]\omega_2\\ +\left[\frac{1}{3}\kappa_2^2 - \frac{1}{5}\kappa_2\kappa_0\right]\omega_4\end{array} & 0 & \begin{array}{c}\left[\frac{1}{5}\kappa_2^2 - \frac{1}{3}\kappa_4\kappa_2\right]\omega_0\\ +\left[\frac{1}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{5}\kappa_2\kappa_0\right]\omega_2\\ +\left[\frac{1}{5}\kappa_0^2 - \frac{1}{3}\kappa_2\kappa_0\right]\omega_4\end{array} \end{pmatrix},$$

which we represent as

$$\mathbf{A}\boldsymbol{T}_2 = \begin{pmatrix} \phi_{110}\omega_0 + \phi_{112}\omega_2 + \phi_{114}\omega_4 & 0 & \phi_{120}\omega_0 + \phi_{122}\omega_2 + \phi_{124}\omega_4 \\ 0 & \phi_{002}\omega_2 & 0 \\ \phi_{210}\omega_0 + \phi_{212}\omega_2 + \phi_{214}\omega_4 & 0 & \phi_{220}\omega_0 + \phi_{222}\omega_2 + \phi_{224}\omega_4 \end{pmatrix},$$

where $\phi_{002} = \pi / \left(3\kappa_2^2\right)$ and

$$\begin{aligned}
\phi_{110} &= \pi\left[\kappa_4^2 - \frac{1}{3}\kappa_4\kappa_2\right], \phi_{112} = \pi\left[\frac{1}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - 2\kappa_4\kappa_2\right], \phi_{114} = \pi\left[\kappa_2^2 - \frac{1}{3}\kappa_2\kappa_0\right], \\
\phi_{120} &= \pi\left[\frac{1}{3}\kappa_2^2 - \kappa_4\kappa_2\right], \phi_{122} = \pi\left[\kappa_4\kappa_0 + \kappa_2^2 - \frac{2}{3}\kappa_2\kappa_0\right], \phi_{124} = \pi\left[\frac{1}{3}\kappa_0^2 - \kappa_2\kappa_0\right], \\
\phi_{210} &= \pi\left[\frac{1}{3}\kappa_4^2 - \frac{1}{5}\kappa_4\kappa_2\right], \phi_{212} = \pi\left[\frac{1}{5}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{3}\kappa_4\kappa_2\right], \phi_{214} = \pi\left[\frac{1}{3}\kappa_2^2 - \frac{1}{5}\kappa_2\kappa_0\right], \\
\phi_{220} &= \pi\left[\frac{1}{5}\kappa_2^2 - \frac{1}{3}\kappa_4\kappa_2\right], \phi_{222} = \pi\left[\frac{1}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{5}\kappa_2\kappa_0\right], \phi_{224} = \pi\left[\frac{1}{5}\kappa_0^2 - \frac{1}{3}\kappa_2\kappa_0\right].
\end{aligned}$$

Note that in this notation

$$\begin{aligned}
\rho_0(m) &= \left\{ \begin{array}{c} \pi\left[\kappa_4^2 - \frac{2}{3}\kappa_4\kappa_2 + \frac{1}{5}\kappa_2^2\right] + \pi\left[\frac{1}{3\kappa_2^2} - 2\kappa_4\kappa_2 + \frac{2}{3}\left(\kappa_4\kappa_0 + \kappa_2^2\right) - \frac{2}{5}\kappa_2\kappa_0\right]\mu_2 \\ +\pi\left[\kappa_2^2 - \frac{2}{3}\kappa_2\kappa_0 + \frac{1}{5}\kappa_0^2\right]\mu_4 \end{array} \right\} \\
&= \left[\phi_{110} + \phi_{220}\right] + \left[\phi_{002} + \phi_{112} + \phi_{222}\right]\mu_2 + \left[\phi_{114} + \phi_{224}\right]\mu_4.
\end{aligned}$$

The characteristic roots of $\mathbf{A}\boldsymbol{T}_2$ are $\rho_1(m) = \phi_{002}\omega_2$ and the two characteristic roots of

$$\begin{pmatrix} \phi_{110}\omega_0 + \phi_{112}\omega_2 + \phi_{114}\omega_4 & \phi_{120}\omega_0 + \phi_{122}\omega_2 + \phi_{124}\omega_4 \\ \phi_{210}\omega_0 + \phi_{212}\omega_2 + \phi_{214}\omega_4 & \phi_{220}\omega_0 + \phi_{222}\omega_2 + \phi_{224}\omega_4 \end{pmatrix} \stackrel{def}{=} \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}.$$

Of these two roots, one is uniformly greater than the other, and is

$$\rho_2(m) = \frac{\psi_{11} + \psi_{22}}{2} + \left\{\left(\frac{\psi_{11} - \psi_{22}}{2}\right)^2 + \psi_{12}\psi_{21}\right\}^{1/2}.$$

Thus the loss is the greater of

$$\begin{aligned}
\mathcal{L}_1(m) &= (1-\nu)\,\rho_0(m) + \nu\rho_1(m), \\
\mathcal{L}_2(m) &= (1-\nu)\,\rho_0(m) + \nu\rho_2(m).
\end{aligned}$$

We apply Theorem 1 of Daemi and Wiens (2013), by which we may proceed as follows. We first find a density m_1 minimizing $\mathcal{L}_1(m)$ in the class of densities for which $\mathcal{L}_1(m) = \max\left(\mathcal{L}_1(m), \mathcal{L}_2(m)\right)$, and a density m_2 minimizing $\mathcal{L}_2(m)$ in the class of densities for which $\mathcal{L}_2(m) = \max\left(\mathcal{L}_1(m), \mathcal{L}_2(m)\right)$. Then the optimal design ξ_* has density

$$m_* = \begin{cases} m_1, & \text{if } \mathcal{L}_1(m_1) \le \mathcal{L}_2(m_2), \\ m_2, & \text{if } \mathcal{L}_2(m_2) \le \mathcal{L}_1(m_1). \end{cases}$$

The two minimizations are first carried out with $\mu_2, \mu_4, \kappa_0, \kappa_2, \kappa_4$ held fixed, thus fixing all ϕ_{ijk} and $\rho_0(m)$. Under these constraints $\mathcal{L}_1(m_1) \leq \mathcal{L}_2(m_2)$ iff $\rho_1(m_1) \leq \rho_2(m_2)$. We first illustrate the calculations for m_2.

We seek

$$\begin{aligned}
&m_2 = \arg\min \rho_2(m), \text{ subject to} \\
&\int_{-1}^{1} m(x)\,dx = 1, \int_{-1}^{1} x^2 m(x)\,dx = \mu_2, \int_{-1}^{1} x^4 m(x)\,dx = \mu_4, \\
&\int_{-1}^{1} \frac{m(x)}{\sigma(x)} dx = \kappa_0, \int_{-1}^{1} x^2 \frac{m(x)}{\sigma(x)} dx = \kappa_2, \int_{-1}^{1} x^4 \frac{m(x)}{\sigma(x)} dx = \kappa_4, \\
&\rho_2(m) - \rho_1(m) - \beta^2 = 0,
\end{aligned}$$

where β is a slack variable. For densities $m(x)$, and with

$$m_{(t)}(x) = (1-t)\,m_1(x) + t m(x),$$

it is sufficient to find m_1 for which the Lagrangian

$$\begin{aligned}
\Phi(t;\boldsymbol{\lambda}) = \rho_2(m_{(t)}) - 2\int_{-1}^{1} &\left\{\left[\lambda_1 + \frac{\lambda_4}{\sigma(x)}\right] + x^2\left[\lambda_2 + \frac{\lambda_5}{\sigma(x)}\right] + x^4\left[\lambda_3 + \frac{\lambda_6}{\sigma(x)}\right]\right\} m_{(t)} dx \\
&- \lambda_7\left(\rho_2(m_{(t)}) - \rho_1(m_{(t)})\right)
\end{aligned}$$

is minimized at $t = 0$ for every $m(\cdot)$ and satisfies the side conditions. The first order condition is

$$\begin{aligned}
0 \leq \Phi'(0;\boldsymbol{\lambda}) &= (1-\lambda_7)\frac{d}{dt}\rho_2\left(m_{(t)}\right)_{|t=0} + \lambda_7 \frac{d}{dt}\rho_1\left(m_{(t)}\right)_{|t=0} \\
&- 2\int_{-1}^{1}\left\{\left[\lambda_1 + \frac{\lambda_4}{\sigma(x)}\right] + x^2\left[\lambda_2 + \frac{\lambda_5}{\sigma(x)}\right] + x^4\left[\lambda_3 + \frac{\lambda_6}{\sigma(x)}\right]\right\}(m(x) - m_2(x))\,dx.
\end{aligned} \tag{B.11}$$

We have that

$$\frac{d}{dt}\rho_1\left(m_{(t)}\right)_{|t=0} = 2\phi_{002}\int_{-1}^{1}\left(\frac{m_2(x)}{\sigma(x)}\right)(m(x) - m_2(x))\,dx,$$

and, with $\rho_2(m)$ represented in an obvious manner as $\rho_2(m) = \psi_0(m) + \sqrt{\psi_1(m)}$,

$$\begin{aligned}
&\frac{d}{dt}\rho_2\left(m_{(t)}\right)_{|t=0} \\
&= \frac{d}{dt}\psi_0\left(m_{(t)}\right)_{|t=0} + \frac{1}{2\sqrt{\psi_1(m_2)}}\frac{d}{dt}\psi_1\left(m_{(t)}\right)_{|t=0} \\
&= \frac{1}{2}\frac{d}{dt}\left\{\int_{-1}^{1}\left\{[\phi_{110} + \phi_{220}] + [\phi_{112} + \phi_{222}]\,x^2 + [\phi_{114} + \phi_{224}]\,x^4\right\}\left(\frac{m_{(t)}(x)}{\sigma(x)}\right)^2 dx\right\}_{|t=0} \\
&+ \frac{1}{2\sqrt{\psi_1(m_2)}}\left\{\begin{array}{c}\left[2\frac{\psi_{11}(m_2) - \psi_{22}(m_2)}{2}\right]\cdot\frac{1}{2}\frac{d}{dt}\left[\psi_{11}\left(m_{(t)}\right) - \psi_{22}\left(m_{(t)}\right)\right] \\ + \frac{d}{dt}\psi_{12}\left(m_{(t)}\right)\psi_{21}(m_2) + \psi_{12}(m_2)\frac{d}{dt}\psi_{21}\left(m_{(t)}\right)\end{array}\right\}_{|t=0},
\end{aligned}$$

which continues as

$$
\begin{aligned}
&\frac{d}{dt}\rho_2\left(m_{(t)}\right)_{|t=0} \\
&= \int_{-1}^{1} \left\{[\phi_{110}+\phi_{220}] + [\phi_{112}+\phi_{222}]\,x^2 + [\phi_{114}+\phi_{224}]\,x^4\right\} \left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx \\
&+ \frac{1}{2\sqrt{\psi_1(m_2)}} \left\{ \begin{array}{c}
[\psi_{11}(m_2)-\psi_{22}(m_2)] \int_{-1}^{1} \left\{ \begin{array}{c} [\phi_{110}-\phi_{220}] \\ +[\phi_{112}-\phi_{222}]\,x^2 \\ +[\phi_{114}-\phi_{224}]\,x^4 \end{array} \right\} \left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx \\
+\psi_{21}(m_2)\int_{-1}^{1}\left\{\phi_{120}+\phi_{122}x^2+\phi_{124}x^4\right\}\left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx \\
+\psi_{12}(m_2)\int_{-1}^{1}\left\{\phi_{210}+\phi_{212}x^2+\phi_{214}x^4\right\}\left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx
\end{array} \right\} \\
&= \int_{-1}^{1}\left(K_0+K_2x^2+K_4x^4\right)\left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx,
\end{aligned}
$$

for

$$
\begin{aligned}
K_0 &= \phi_{110}+\phi_{220}+\frac{[\psi_{11}(m_2)-\psi_{22}(m_2)]\,[\phi_{110}-\phi_{220}]+\psi_{21}(m_2)\,\phi_{120}+\psi_{12}(m_2)\,\phi_{210}}{2\sqrt{\psi_1(m_2)}}, \\
K_2 &= \phi_{112}+\phi_{222}+\frac{[\psi_{11}(m_2)-\psi_{22}(m_2)]\,[\phi_{112}-\phi_{222}]+\psi_{21}(m_2)\,\phi_{122}+\psi_{12}(m_2)\,\phi_{212}}{2\sqrt{\psi_1(m_2)}}, \\
K_4 &= \phi_{114}+\phi_{224}+\frac{[\psi_{11}(m_2)-\psi_{22}(m_2)]\,[\phi_{114}-\phi_{224}]+\psi_{21}(m_2)\,\phi_{124}+\psi_{12}(m_2)\,\phi_{214}}{2\sqrt{\psi_1(m_2)}}.
\end{aligned}
$$

Substituting into (B.11) gives

$$
\begin{aligned}
\Phi'(0;\boldsymbol{\lambda}) &= (1-\lambda_7)\int_{-1}^{1}\left(K_0+K_2x^2+K_4x^4\right)\left(\frac{m_2(x)}{\sigma^2(x)}\right)(m(x)-m_2(x))\,dx \\
&+ \lambda_7\cdot 2\phi_{002}\int_{-1}^{1}\left(\frac{m_2(x)}{\sigma(x)}\right)(m(x)-m_2(x))\,dx \\
&- 2\int_{-1}^{1}\left\{\left[\lambda_1+\frac{\lambda_4}{\sigma(x)}\right]+x^2\left[\lambda_2+\frac{\lambda_5}{\sigma(x)}\right]+x^4\left[\lambda_3+\frac{\lambda_6}{\sigma(x)}\right]\right\}(m(x)-m_2(x))\,dx \\
&= \int_{-1}^{1}\left\{ \begin{array}{c} \left\{\left[(1-\lambda_7)\left(\frac{K_0+K_2x^2+K_4x^4}{\sigma^2(x)}\right)+\frac{2\lambda_7\phi_{002}}{\sigma(x)}\right]m_2(x)\right\} \\ -2\left\{\left[\lambda_1+\frac{\lambda_4}{\sigma(x)}\right]+x^2\left[\lambda_2+\frac{\lambda_5}{\sigma(x)}\right]+x^4\left[\lambda_3+\frac{\lambda_6}{\sigma(x)}\right]\right\} \end{array}\right\}(m(x)-m_2(x))\,dx,
\end{aligned}
$$

entailing

$$
m_2(x) = \left(\frac{2\left\{\left[\lambda_1+\frac{\lambda_4}{\sigma(x)}\right]+x^2\left[\lambda_2+\frac{\lambda_5}{\sigma(x)}\right]+x^4\left[\lambda_3+\frac{\lambda_6}{\sigma(x)}\right]\right\}}{\left[(1-\lambda_7)\left(\frac{K_0+K_2x^2+K_4x^4}{\sigma^2(x)}\right)+\frac{2\lambda_7\phi_{002}}{\sigma(x)}\right]}\right)^{+}.
$$

The derivation of m_1 is very similar. We seek

$$\begin{aligned}
&m_1 = \arg\min \rho_1(m), \text{ subject to} \\
&\int_{-1}^{1} m(x)\,dx = 1, \int_{-1}^{1} x^2 m(x)\,dx = \mu_2, \int_{-1}^{1} x^4 m(x)\,dx = \mu_4, \\
&\int_{-1}^{1} \frac{m(x)}{\sigma(x)}dx = \kappa_0, \int_{-1}^{1} x^2\frac{m(x)}{\sigma(x)}dx = \kappa_2, \int_{-1}^{1} x^4\frac{m(x)}{\sigma(x)}dx = \kappa_4, \\
&\rho_1(m) - \rho_2(m) - \beta^2 = 0,
\end{aligned}$$

where β is a slack variable. For densities $m(x)$, and with

$$m_{(t)}(x) = (1-t)\,m_2(x) + tm(x),$$

it is sufficient to find m_2 for which the Lagrangian

$$\begin{aligned}
\Phi(t;\boldsymbol{\lambda}) = \rho_1(m_{(t)}) - 2\int_{-1}^{1}\left\{\left[\lambda_1 + \frac{\lambda_4}{\sigma(x)}\right] + x^2\left[\lambda_2 + \frac{\lambda_5}{\sigma(x)}\right] + x^4\left[\lambda_3 + \frac{\lambda_6}{\sigma(x)}\right]\right\} m_{(t)}dx \\
- \lambda_7\left(\rho_1(m_{(t)}) - \rho_2(m_{(t)})\right)
\end{aligned}$$

is minimized at $t = 0$ for every $m(\cdot)$ and satisfies the side conditions. This leads to the same first order condition as (B.11), except that λ_7 is replaced by $1 - \lambda_7$; this in turn leads to

$$m_1(x) = \left(\frac{2\left\{\left[\lambda_1 + \frac{\lambda_4}{\sigma(x)}\right] + x^2\left[\lambda_2 + \frac{\lambda_5}{\sigma(x)}\right] + x^4\left[\lambda_3 + \frac{\lambda_6}{\sigma(x)}\right]\right\}}{\left[\lambda_7\left(\frac{K_0 + K_2x^2 + K_4x^4}{\sigma^2(x)}\right) + \frac{2(1-\lambda_7)\phi_{002}}{\sigma(x)}\right]}\right)^+.$$

In either case, the minimizing design density is of the form

$$m(x;\mathbf{a}) = \left(\frac{q_1(x) + \frac{q_2(x)}{\sigma(x)}}{\frac{a_{00}}{\sigma(x)} + \frac{q_3(x)}{\sigma^2(x)}}\right)^+, \tag{B.12}$$

for polynomials $q_j(x) = a_{0j} + a_{2j}x^2 + a_{4j}x^4$, $j = 1, 2, 3$. The constants a_{ij} forming $\mathbf{a}$ are determined by the constraints in terms of the μ_k, κ_k and β^2, which are then optimally chosen to minimize the loss. It is simpler however to choose $\mathbf{a}$ directly, to minimize $\mathcal{L}_\nu(\xi|\sigma)$ over all densities of the form (B.12), subject to $\int_{-1}^{1} m(x;\mathbf{a})dx = 1$.

Acknowledgements

This work has been supported by the Natural Sciences and Engineering Research Council of Canada.

References

Daemi, M., and Wiens, D. P. (2013), "Techniques for the Construction of Robust Regression Designs," *The Canadian Journal of Statistics*, 41, 679 - 695.

Knight, K. (1998), "Limiting Distributions for l_1 Estimators Under General Conditions," *Annals of Statistics*, 26, 755-770.

Pollard, D. (1991), "Asymptotics for Least Absolute Deviation Regression Estimators," *Econometric Theory*, 7, 186-199.