The original constrained problem

The original constrained problem is stated as

$$\min_{\tau} \iint -\operatorname{sgn}(v)\, u\, f_Y(u, v; \tau)\; du\, dv$$
$$\text{subject to } \kappa - \iint \operatorname{sgn}(v)\, w\, f_Z(w, v; \tau)\; dw\, dv \geq 0, \text{and } \tau^{\intercal}\tau = 1.$$

Suppose that the strict feasible set strict($\mathcal{F}$) is non-empty, and let τ^0 denote a constrained minimizer of this original problem. For simplicity, we let $g(\tau) = -\iint \operatorname{sgn}(v) u\, f_Y(u, v; \tau)\, du\, dv$, $c_1(\tau) = \kappa - \iint \operatorname{sgn}(v) w f_Z(w, v; \tau)\, dw\, dv$, and $c_2(\tau) = \tau^{\intercal}\tau$. The value of $g(\tau)$ at $\tau = \tau^0$, $g(\tau^0)$, is denoted by g^0. Similarly, c_i^0 denotes the value of $c_i(\tau)$ at $\tau = \tau^0$, $c_i(\tau^0)$, for $i = 1, 2$. Also, $\mathcal{A}^0$ denotes the set of active constraint at τ^0, $\mathcal{A}(\tau^0)$. In our current case, it is either $\mathcal{A}^0 = \{c_1^0,\, c_2^0\}$, or $\mathcal{A}^0 = \{c_2^0\}$.

Perturbed KKT conditions

Fmincon interior point algorithm can be interpreted as log barrier penalty with slack variables, i.e.,

$$\min_{\tau,\, s} B(\tau, s) = \min_{\tau,\, s} \iint -\operatorname{sgn}(v) u f_Y(u, v; \tau)\, du\, dv - \mu \log s$$
$$\text{subject to } \kappa - \iint \operatorname{sgn}(v)\, w\, f_Z(w, v; \tau)\; dw\, dv - s = 0 \text{ and } \tau^{\intercal}\tau - 1 = 0,$$

where μ is a sequence of positive decreasing small constants converging to zero. Its solution does not coincide with that of the original constrained problem for $\mu > 0$. The barrier approach consists of finding (approximate) solutions of the barrier problem for a sequence of positive barrier parameters μ_k that converges to zero.

KKT conditions for constrained problem

Problem:

$$\min_{x} f(x)$$
$$\begin{aligned} \text{subject to } & c_E(x) = 0 \\ & c_I(x) - s = 0 \\ & s \geq 0 \end{aligned}$$

Log-barrier method problem with slack variables S:

$$\min_{x,s} f\left(x\right) - \lambda \sum_{i=1}^{m} \log s_i$$
$$c_E\left(x\right) = 0$$
$$c_I\left(x\right) - s = 0$$

Note: $s_i > 0$ is satisfied automatically due to logarithm.

Perturbed KKT conditions for log-barrier:

$$\begin{aligned}
&\nabla f\left(x\right) - A_E^T\left(x\right) y - A_I^T\left(x\right) z = 0 \\
&-\lambda S^{-1} e + z = 0 \text{ or } -\lambda e + Sz = 0 \\
&c_E\left(x\right) = 0 \\
&c_I\left(x\right) - s = 0
\end{aligned}$$

Note: $\lambda e = -Sz$

My problem:

$$\min_{\tau} \iint -sgn\left(v\right) u\, f_Y\left(u, v; \tau\right)\, du\, dv$$
$$\text{subject to } \kappa - \iint \operatorname{sgn}\left(v\right) w\, f_Z\left(w, v; \tau\right)\, dw\, dv \geq 0$$

Log-barrier formation:

$$\min_{\tau,s} \iint -sgn\left(\nu\right) u f_Y\left(u, \nu; \tau\right)\, du\, dv - \lambda \log s_1$$
$$\text{subject to } \kappa - \iint \operatorname{sgn}\left(\nu\right) w f_z\left(w, v; \tau\right)\, dw\, dv - s_1 = 0$$

Note: $s_1 > 0$ is satisfied automatically due to logarithm again.

Perturbed KKT conditions:

$\exists\lambda, z_1, s_1$, such that

$$\iint -sgn\left(\nu\right)u\nabla_{\tau}f_Y\left(u,\nu;\tau\right)dud\nu - z_1\left\{\kappa - \iint sgn\left(\nu\right)\omega\nabla_{\tau}f_Z\left(\omega,\nu;\tau\right)dwd\nu\right\} = 0$$

$$\kappa - \iint sgn\left(\nu\right)w\nabla_{\tau}f_Z\left(\omega,\nu;\tau\right)dwd\nu - s_1 = 0$$

$$-\lambda/s_1 + z_1 = 0$$

From the last equation, we have $s_1 = \lambda/z_1 > 0$
There is theorem which gives the conditions under which, for sufficiently small μ, the sequence $\{\tau_{\mu}^*\}$ defines a differentiable penalty-barrier trajectory converging to τ_{μ}^0.

To find τ_{μ}^*, we exploit its stationarity. The gradient of $\Phi_{PB}(\tau,\mu)$ is

$$\begin{aligned}\nabla_{\tau}\Phi_{PB}(\tau,\mu) = &\iint -\mathrm{sgn}(v)u\nabla_{\tau}f_Y(u,v;\tau)\,du\,dv\\ &+\mu\,\frac{\iint \mathrm{sgn}(v)w\nabla_{\tau}f_Z(w,v;\tau)\,dw\,dv}{\kappa-\iint \mathrm{sgn}(v)wf_Z(w,v;\tau)\,dw\,dv}+\frac{2}{\mu}\,(\tau^{\mathsf{T}}\tau-1)\tau,\end{aligned}$$

noting that ∇_{τ} represents the first order derivative with respect to τ. If we are willing to assume that $\Phi_{PB}(\tau,\mu)$ is twice-continuously differentiable, it must hold that $\nabla\Phi_{PB}(\tau_{\mu}^*,\mu) = 0$ to satisfy the stationarity, i.e.,

$$\iint \mathrm{sgn}\left(v\right)u\,\nabla_{\tau}f_Y\left(u,v;\tau_{\mu}^*\right)\,du\,dv = \mu\,\frac{\iint \mathrm{sgn}(v)w\nabla_{\tau}f_Z(w,v;\tau_{\mu}^*)\,dw\,dv}{\kappa-\iint \mathrm{sgn}(v)\,w\,f_Z(w,v;\tau_{\mu}^*)\,dw\,dv}+\frac{2}{\mu}\,(\tau_{\mu}^{*\mathsf{T}}\tau_{\mu}^*-1)\tau_{\mu}^*,$$

with $\tau_{\mu}^{*\mathsf{T}}\tau_{\mu}^{*\mathsf{T}}-1=0$. The barrier multiplier, the coefficient in this linear relationship above, denoted by λ_{μ}, is defined as

$$\lambda_{\mu} \triangleq \frac{\mu}{\kappa-\iint \mathrm{sgn}(v)wf_Z(w,v;\tau_{\mu})\,dw\,dv}.$$

This relationship can be re-written as

$$\lambda_{\mu}\left[\kappa-\iint \mathrm{sgn}(v)wf_Z(w,v;\tau_{\mu})\,dw\,dv\right]=\mu.$$

This relationship between the barrier multiplier, the constraint value, and the barrier parameter, called perturbed complementarity, is analogous as $\mu\to 0$ to the complementarity condition $c(\tau^*)\lambda^*=0$ that holds at a KKT point.

Estimation of the log-barrier penalty function
To estimate the log-barrier penalty function, we use kernel density estimators, denoted by

$\widehat{f}_Y(u,v;\tau)$ and $\widehat{f}_Z(w,v;\tau)$, to estimate the corresponding density functions. Hence, the estimated log-barrier function is

$$\widehat{B}(\tau,\mu) = \iint -\mathrm{sgn}(v)u\widehat{f}_Y(u,v;\tau)\,du\,dv - \mu\ln\left[\kappa - \iint \mathrm{sgn}(v)w\widehat{f}_Z(w,v;\tau)\,dw\,dv\right],$$

and the gradient of the estimator is

$$\begin{aligned}\nabla\widehat{B}(\tau,\mu) &= \iint -\mathrm{sgn}(v)u\nabla\widehat{f}_Y(u,v;\tau)\,du\,dv + \mu\frac{\iint \mathrm{sgn}(v)w\nabla\widehat{f}_Z(w,v;\tau)\,dw\,dv}{\kappa - \iint \mathrm{sgn}(v)w\widehat{f}_Z(w,v;\tau)\,dw\,dv}\\ &= \iint -\mathrm{sgn}(v)u\nabla\widehat{f}_Y(u,v;\tau)\,du\,dv + \widehat{\lambda}_\mu \iint \mathrm{sgn}(v)w\nabla\widehat{f}_Z(w,v;\tau)\,dw\,dv,\end{aligned}$$

where $\widehat{\lambda}_\mu(\tau) = {}^{\mu}/_{\kappa - \iint \mathrm{sgn}(v)w\widehat{f}_Z(w,v;\tau)\,dw\,dv}$.

Consistency of $\widehat{\tau}^k$ and $\widehat{\lambda}_\mu$.
We need to prove that $\widehat{\tau}^k$ is a consistent estimator of τ^{*k}.
$\widehat{\tau}^k - \tau^{*k} = O_p(n^{1/2})$, and $\widehat{\lambda}^k - \lambda^{*k} = O_p(n^{1/2})$.
Theorem proved that λ_μ is bounded.

Asymptotic distribution of $\widehat{\tau}^k$
Estimating equations:

$$\nabla\widehat{B}(\tau,\mu) = \iint -\mathrm{sgn}(v)u\nabla\widehat{f}_Y(u,v;\tau)\,du\,dv + \hat{\lambda}_\mu(\tau) \iint \mathrm{sgn}(v)w\nabla\widehat{f}_Z(w,v;\tau)\,dw\,dv = 0$$

where $\hat{\lambda}_\mu(\tau) = {}^{\mu}/_{[\kappa - \iint \mathrm{sgn}(v)w\widehat{f}_Z(w,v;\tau)\,dw\,dv]}$.

$$\begin{aligned}\nabla\widehat{B}(\tau,\mu) &= \iint -\mathrm{sgn}(v)u\nabla\widehat{f}_Y(u,v;\tau)\,du\,dv + \hat{\lambda}_\mu \iint \mathrm{sgn}(v)w\nabla\widehat{f}_Z(w,v;\tau)\,dw\,dv\\ &= -\frac{2}{n}\sum_{i=1}^{n}\boldsymbol{X}_{i,1}^\intercal\boldsymbol{\beta}_{Y1}k\left(-\frac{\boldsymbol{X}_i^\intercal\boldsymbol{\tau}}{h}\right)\boldsymbol{X}_i + \hat{\lambda}_\mu(\tau)\frac{2}{n}\sum_{i=1}^{n}\boldsymbol{X}_{i,1}^\intercal\boldsymbol{\beta}_{Z1}k\left(-\frac{\boldsymbol{X}_i^\intercal\boldsymbol{\tau}}{h}\right)\boldsymbol{X}_i\\ &= N(\mu_1,\Sigma_1) + C_pN(\mu_2,\Sigma_2)\end{aligned}$$

$$
\begin{aligned}
\nabla^2 \widehat{B}(\tau,\mu) =& \iint -\mathrm{sgn}(v) u \nabla^2 \widehat{f}_Y(u,v;\tau)\, du\, dv + \\
& \nabla \hat{\lambda}_\mu(\tau) \iint \mathrm{sgn}(v) w \nabla \widehat{f}_Z(w,v;\tau)\, dw\, dv + \hat{\lambda}_\mu(\tau) \iint \mathrm{sgn}(v) w \nabla^2 \widehat{f}_Z(w,v;\tau)\, dw\, dv \\
=& -\frac{2}{nh} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^\intercal \left(\hat{\lambda}_\mu(\tau) \boldsymbol{\beta}_{Z1} - \boldsymbol{\beta}_{Y1} \right) k' \left(-\frac{\boldsymbol{X}_i^\intercal \boldsymbol{\tau}}{h} \right) \boldsymbol{X}_i \boldsymbol{X}_i^\intercal + \\
& \frac{2}{n} \sum_{i=1}^{n} \nabla \hat{\lambda}_\mu(\tau) \boldsymbol{X}_{i,1}^\intercal \boldsymbol{\beta}_{Z1} k \left(-\frac{\boldsymbol{X}_i^\intercal \boldsymbol{\tau}}{h} \right) \boldsymbol{X}_i
\end{aligned}
$$

$$
\begin{aligned}
\nabla \widehat{\lambda}_\mu(\tau) =& \frac{\mu}{\left(\kappa - \iint \mathrm{sgn}(v) w \widehat{f}_Z(w,v;\tau)\, dw\, dv \right)^2} \iint \mathrm{sgn}(v) w \nabla \widehat{f}_Z(w,v;\tau)\, dw\, dv \\
=& \mu \left[\kappa - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^\intercal \boldsymbol{\beta}_{Z1} \left\{ 1 - 2K \left(-\frac{\boldsymbol{X}_i^\intercal \boldsymbol{\tau}}{h} \right) \right\} \right]^{-1} \left[\frac{2}{n} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^\intercal \boldsymbol{\beta}_{Z1} k \left(-\frac{\boldsymbol{X}_i^\intercal \boldsymbol{\tau}}{h} \right) \boldsymbol{X}_i \right]
\end{aligned}
$$

[Notation: k and κ looks to similar]
[Need to estimate $\hat{\beta}$ too]
Need to prove that the difference between $\widehat{B}_n(\tau, \hat{\beta}, \mu)$ and $\widehat{B}_n(\tau, \beta^*, \mu)$ is negligible? i.e., $\widehat{B}_n(\beta^*) - \widehat{B}_n(\hat{\beta}) = O_p(n^{-1/2})$

Taylor expansion of $\nabla \widehat{B}(\tau^{*k}, \mu)$ at $\tau = \hat{\tau}^k$ shows that

$$
\nabla \widehat{B}(\tau^{*k}, \mu) = \nabla \widehat{B}(\hat{\tau}^k, \mu) - \nabla^2 \widehat{B}(\tilde{\tau}^k, \mu)(\widehat{\tau}^k - \tau^{*k}),
$$

where $\tilde{\tau}^k$ is between τ^{*k} and $\widehat{\tau}^k$. As $\hat{\tau}^k$ is the minimizer of $B(\tau, \mu)$, it satisfies the first order condition that $\nabla B(\hat{\tau}^k, \mu) = 0$. Therefore, we have

$$
\sqrt{n} \nabla \widehat{B}(\tau^{*k}, \mu) = -\sqrt{n} \nabla^2 \widehat{B}(\tilde{\tau}^k, \mu)(\widehat{\tau}^k - \tau^{*k}).
$$

Derivation of the integrations
The integration we need

$$
\begin{aligned}
& \iint \mathrm{sgn}\,(v)\, u\, f\,(u, v; \tau, \beta_{\cdot 1})\; du\, dv \\
=& 2 \iint u\, \mathbb{I}\,(v \geq 0)\, f\,(v, u; \tau, \beta_{\cdot 1})\; dv\, du - \int u\, f\,(u; \beta_{\cdot 1})\; du
\end{aligned}
$$

The estimator is

$$
\begin{aligned}
&\iint \operatorname{sgn}(v)\, u\, \widehat{f}_n\left(u, v; \tau, \beta_{\cdot 1}\right)\, du\, dv \\
=&2 \iint u\, \mathbb{I}\left(v \geq 0\right) \widehat{f}_n\left(v, u; \tau, \beta_{\cdot 1}\right)\, dv\, du - \int u\, \widehat{f}_n\left(u; \tau, \beta_{\cdot 1}\right)\, du \\
=&\frac{2}{nh^2} \iint u\, \mathbb{I}\left(v \geq 0\right) \sum_{i=1}^{n} k\left(\frac{v - V_i}{h}\right) k\left(\frac{u - U_i}{h}\right)\, du\, dv - \\
&\frac{1}{nh} \int u \sum_{i=1}^{n} k\left(\frac{u - U_i}{h}\right)\, du \\
=&\frac{2}{n} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^{\intercal} \boldsymbol{\beta}_{\cdot 1} \left\{1 - K\left(-\frac{\boldsymbol{X}_i^{\intercal} \boldsymbol{\tau}}{h}\right)\right\} - \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^{\intercal} \boldsymbol{\beta}_{\cdot 1} \\
=&\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_{i,1}^{\intercal} \boldsymbol{\beta}_{\cdot 1} \left\{1 - 2K\left(-\frac{\boldsymbol{X}_i^{\intercal} \boldsymbol{\tau}}{h}\right)\right\}
\end{aligned}
$$

where $\widehat{f}_n(u_1, u_2; \tau, \widehat{\beta}_{\cdot 1})$ are the kernel density estimator for $(X^{\intercal}\tau, X^{\intercal}\beta_{\cdot 1})$ with the forms of

$$
\widehat{f}_n(u, v; \boldsymbol{\tau}, \widehat{\boldsymbol{\beta}}_{\cdot 1}) = \frac{1}{nh^2} \sum_{i=1}^{n} k\left(\frac{u - U_i}{h}\right) k\left(\frac{v - V_i}{h}\right).
$$

Moreover, $K(s)$ is the corresponding CDF of the kernel function $k(s)$, which is chosen to be a symmetric probability density. More precisely, $k(s)$ satisfies the following assumptions:

1. $\int_{-\infty}^{\infty} k(s)\, ds = 1$.
2. $k(s) > 0$ for all s.
3. $k(-s) = k(s)$ for all s.
4. The first order derivative of the kernel, $k'(s)$, exists and is bounded.

The last equality above holds by following the derivation.

We first derive $\frac{2}{h^2} \iint u_2\, \mathbb{I}\left(u_1 \geq 0\right) k\left(\frac{u_1 - U_{i,1}}{h}\right) k\left(\frac{u_2 - U_{i,2}}{h}\right)\, du_1\, du_2$. Let $s = \frac{u_1 - U_{i,1}}{h}$ and $t =$

$\frac{u_2-U_{i,2}}{h}$. Then, $u_1 = U_{i,1} + sh$ and $u_2 = U_{i,2} + th$. Also, $du_1 = h\,ds$ and $du_2 = h\,dt$.

$$
\begin{aligned}
&\frac{2}{h^2}\iint u_2\,\mathbb{I}\left(u_1 \geq 0\right) k\left(\frac{u_1 - U_{i,1}}{h}\right) k\left(\frac{u_2 - U_{i,2}}{h}\right)\,du_1\,du_2\\
=&2\iint \left(U_{i,2} + th\right)\mathbb{I}\left(U_{i,1} + sh \geq 0\right) k\left(s\right) k\left(t\right)\,ds\,dt\\
=&2\int U_{i,2}\,\mathbb{I}\left(s \geq -\frac{U_{i,1}}{h}\right) k\left(s\right)\,ds\\
=&2U_{i,2}\left\{1 - K\left(-\frac{U_{i,1}}{h}\right)\right\}\\
=&2\boldsymbol{X}_{i,1}^{\intercal}\boldsymbol{\beta}_{\cdot 1}\left\{1 - K\left(-\frac{\boldsymbol{X}_i^{\intercal}\boldsymbol{\tau}}{h}\right)\right\},
\end{aligned}
$$

where $K\left(s\right) = \int k\left(s\right)\,ds + c$. The second equality holds, as $\int k(t)\,dt = 1$ and $\int t\,k(t)\,dt = 0$. The third equality holds as $\int \mathbb{I}\left(s \geq -\frac{U_{i,1}}{h}\right) k\left(s\right)\,ds = 1 - \int_{-\infty}^{-U_{i,1}/h} k\left(s\right)\,ds = 1 - K\left(-\frac{U_{i,1}}{h}\right)$, where $U_{i,1} = \boldsymbol{X}_i^{\intercal}\boldsymbol{\tau}$.

Then, we derive $\frac{1}{h}\int u_2 k(\frac{u_2-U_{i,2}}{h})\,du_2$ by changing variable similarly. Let $t = \frac{u_2-U_{i,2}}{h}$,and we get $u_2 = U_{i,2} + th$, and $du_2 = h\,dt$.

$$
\begin{aligned}
&\frac{1}{h}\int u_2 k\left(\frac{u_2 - U_{i,2}}{h}\right)\,du_2\\
=&\int \left(U_{i,2} + th\right) k\left(t\right)\,dt\\
=&U_{i,2} = \boldsymbol{X}_{i,1}^{\intercal}\boldsymbol{\beta}_{\cdot 1}.
\end{aligned}
$$

Again, the second equality holds as $\int k(t)\,dt = 1$, and $\int t\,k(t)\,dt = 0$.
The integration over the first-order derivative

$$
\begin{aligned}
&\iint \operatorname{sgn}\left(v\right) u\,\nabla\widehat{f}_n\left(u, v; \tau, \beta_{\cdot 1}\right)\,du\,dv\\
=&\frac{\partial}{\partial\tau}\iint \operatorname{sgn}\left(v\right) u\,\widehat{f}_n\left(u, v; \tau, \beta_{\cdot 1}\right)\,du\,dv\\
=&\frac{2}{n}\sum_{i=1}^{n}\boldsymbol{X}_{i,1}^{\intercal}\boldsymbol{\beta}_{\cdot 1} k\left(-\frac{\boldsymbol{X}_i^{\intercal}\boldsymbol{\tau}}{h}\right)\boldsymbol{X}_i
\end{aligned}
$$