203C HW1

Christopher Ackerman

April 23, 2021

1 Question 1

1. No, this vector isn't identified. In order to identify the entire vector we need to identify each element. We can't identify h within H because we don't see observations on the entire domain $\mathbb{R}$, so we can't identify the vector.

2. Nothing is identified in this case. Denote true values with a $*$, and consider

$$\begin{aligned}
\tilde{h}(x) &= ch^*(x) \\
\tilde{\mu} &= c\mu^* \\
\tilde{\sigma}^2 &= c^2\sigma^{*2} \\
c &> 0 \\
\mathbb{P}(Y = 1 \mid X = x; \tilde{\mu}\tilde{\sigma}, \tilde{h}) &= \Phi\left(\frac{\tilde{h}(x) - \tilde{\mu}}{\sqrt{\tilde{\sigma}^2}}\right) \\
&= \Phi\left(\frac{h^*(x) - \mu^*}{\sqrt{\sigma^{*2}}}\right) \\
&= \mathbb{P}(Y = 1 \mid X = x; \mu^*, \sigma^{*2}, h^*)
\end{aligned}$$

3. With this normalization, we can write

$$\begin{aligned}
\mathbb{P}(Y = 1 \mid X = x) &= \Phi(h(x)) \\
\implies h(x) &= \Phi^{-1}(\mathbb{P}(Y = 1 \mid X = x))
\end{aligned}$$

   Now $h(x)$ is identified in H_x, but it is not identified in H because there are infinitely many points in $\mathbb{R}$ (e.g. $x = 6$) where the value of $h(x)$ could be different.

4. Plugging in this functional form to the result in the last part,

$$\alpha + \beta x = \Phi^{-1}(\mathbb{P}(Y = 1 \mid X = x))$$

   We can pin down values for α and β as

$$\begin{aligned}
\alpha &= \mathbb{E}(Y \mid X = 0) \\
\beta &= \mathbb{E}(Y \mid X = 1) - \mathbb{E}(Y \mid X = 0)
\end{aligned}$$

   Since $x = 0$ and $x = 1$ are in the support of both H and H_X, we have identification within both sets.

2 Question 2

1.

$$\begin{aligned}\mathbb{P}(Y=1 \mid X=x) &= \mathbb{P}(\varepsilon \leq x'\beta \mid X=x) \\ &= \mathbb{P}\left(z \leq \frac{x'\beta-\mu}{\sqrt{s^2(x)}} \mid X=x\right) \\ &= \Phi\left(\frac{x'\beta-\mu}{\sqrt{s^2(x)}}\right)\end{aligned}$$

2. No, β is not identified. Take

$$\begin{aligned}\tilde{\beta} &\equiv c\beta^* \\ \tilde{\mu} &\equiv c\mu^* \\ \tilde{s}^2(x) &\equiv c^2 s^{*2}(x) \\ \mathbb{P}(Y=1 \mid X=x; \tilde{\beta}, \tilde{\mu}, \tilde{s}) &= \Phi\left(\frac{x'\tilde{\beta}-\tilde{\mu}}{\sqrt{\tilde{s}^2(x)}}\right) \\ &= \Phi\left(\frac{x'\beta^*-\mu^*}{\sqrt{s^{*2}(x)}}\right) \\ &= \mathbb{P}(Y=1 \mid X=x; \beta^*, \mu^*, s^{*2})\end{aligned}$$

3. We can now write the conditional expectation as

$$\mathbb{P}(Y=1 \mid x=x) = \Phi\left(\frac{x_1+x_2\beta-\mu}{s^2(x_3)}\right)$$

   We can identify β_2 by differentiation with respect to x_2, and we can identify $\sqrt{s^2(x_3)} = |s(x_3)$ by differentiating with respect to x_1. So β_2 is identified, and s is identified if we're willing to assume that $s(x_3) \geq 0$.

4. Yes; we still have identification for both s and β_2 since the argument in (c) does not depend on the value of $s(0)$.

5. They are both identified. To see that F_ε is identified, define some $\tilde{F}_\varepsilon$ such that

$$F_\varepsilon^*(x_1+v^*(x_2)) = \tilde{F}_\varepsilon(x_1+\tilde{v}(x_2)) \quad \forall x$$

   If we look at $x_2=0$,

$$F_\varepsilon^*(x_1) = \tilde{F}_\varepsilon(x_1),$$

   so F_ε is identified. Now if we look at $x_1=0$,

$$F_\varepsilon(v^*(x_2)) = F_\varepsilon(\tilde{v}(x_2)) \quad \forall x$$

   F_ε is strictly increasing so we can take its inverse, and therefore v is identified.

6. The answer is still the same since we have the normalization/known value $h(\overline{x} = \alpha$. We would only have to change the argument if we didn't know this value.

3 Question 3

1.

$$\begin{aligned}\mathbb{E}(Y \mid X = x) &= \mathbb{P}(Y = 1 \mid X = x) - (1 - \mathbb{P}(Y = 1 \mid X = x)) \\ &= 2\mathbb{P}(Y = 1 \mid X = x) - 1\end{aligned}$$

2.

$$\begin{aligned}\mathbb{P}(Y = 1 \mid X = x) &= \mathbb{P}(\varepsilon \leq v(x_1) - v(x_2) \mid X = x) \\ &= F_{\varepsilon|X=x}(v(x_1) - v(x_2)) \\ v(x_1) - v(x_2) &= t > 0 \implies \\ \mathbb{P}(Y = 1 \mid X = x) &= F_{\varepsilon|X=x}(t) \\ &> F_{\varepsilon|X=x}(0) \\ &= \frac{1}{2} \\ v(x_1) - v(x_2) &= t < 0 \implies \\ \mathbb{P}(Y = 1 \mid X = x) &= F_{\varepsilon|X=x}(t) \\ &< F_{\varepsilon|X=x}(0) \\ &= \frac{1}{2} \\ \therefore \mathbb{P}(Y = 1 \mid X_1 = x_1, X_2 = x_2) &\geq \frac{1}{2} \iff v(x_1) \geq v(x_2)\end{aligned}$$

3. We can define level sets at $V(\overline{x})$ as

$$\begin{aligned}S(V)(\overline{x}) &= \{x \mid V(x) \geq V(\overline{x})\} \\ &= \left\{x \mid \mathbb{P}(Y = 1 \mid X_1 = \overline{x}, X_2 = x) \geq \frac{1}{2}\right\} \\ S'(V)(\overline{x}) &= \{x \mid V(x) \leq V(\overline{x})\} \\ &= \left\{x \mid \mathbb{P}(Y = 1 \mid X_1 = x, X_2 = \overline{x}) \leq \frac{1}{2}\right\}\end{aligned}$$

V is identified when no function in S or S' is an increasing function of another function in that set, and when $V(\overline{x}) = 0$ for some $\overline{x}$.

4. We can only identify the value of v at a single point $t = v(x_1) - v(x_2)$, so $F_{\varepsilon|X=x}(t)$ is not identified.