Math 6417 Homework 2

Jacob Hauck

October 4, 2023

A continuous function $\sigma : \mathbf{R} \to \mathbf{R}$ is called **sigmoidal** if there exists $T > 0$ such that

$$\sigma(t) = \begin{cases} 1 & t \geq T, \\ 0 & t \leq -T. \end{cases} \tag{1}$$

Let σ be sigmoidal in the following problems.

Question 1.

Let $y \in \mathbf{R}^n$, and $\theta, \phi \in \mathbf{R}$. For $x \in \mathbf{R}^n$, define

$$\sigma_\lambda(x; \theta, \phi) = \sigma\left(\lambda\left(y^T x + \theta\right) + \phi\right). \tag{2}$$

Then

$$\sigma_\lambda(x; \theta, \phi) \to \gamma(x) = \begin{cases} 1 & y^T x + \theta > 0 \\ 0 & y^T x + \theta < 0 \\ \sigma(\phi) & y^T x + \theta = 0 \end{cases} \qquad \text{as } \lambda \to \infty. \tag{3}$$

Proof. We use proof by cases.

1. If $y^T x + \theta > 0$, then
$$\lambda \geq \frac{T - \phi}{y^T x + \theta} \implies \lambda(y^T x + \theta) + \phi \geq T \implies \sigma_\lambda(x; \theta, \phi) = 1, \tag{4}$$
so $\sigma_\lambda(x; \theta, \phi) \to 1 = \gamma(x)$ as $\lambda \to \infty$.
2. If $y^T x + \theta < 0$, then
$$\lambda \geq \frac{-T - \phi}{y^T x + \theta} \implies \lambda(y^T x + \theta) + \phi \leq -T \implies \sigma_\lambda(x; \theta, \phi) = 0, \tag{5}$$
so $\sigma_\lambda(x; \theta, \phi) \to 0 = \gamma(x)$ as $\lambda \to \infty$.
3. If $y^T x + \theta = 0$, then $\sigma_\lambda(x; \phi, \theta) = \sigma(\phi)$ for all λ, so $\sigma_\lambda(x; \phi, \theta) \to \sigma(\phi) = \gamma(x)$ as $\lambda \to \infty$.

$\square$

Question 2.

Let $y \in \mathbf{R}^n$, let $\Pi_{y,\theta} = \left\{x \mid y^T x + \theta = 0\right\}$, and let $H_{y,\theta} = \left\{x \mid y^T x + \theta > 0\right\}$. If μ is a finite Borel measure on $[0,1]^n$ such that

$$\int_{[0,1]^n} \sigma_\lambda(x)\, \mathrm{d}\mu(x) = 0 \qquad \text{for all } \lambda, \theta, \phi \in \mathbf{R}, \tag{6}$$

then

$$\sigma(\phi)\mu(\Pi_{y,\theta}) + \mu(H_{y,\theta}) = 0 \qquad \text{for all } \lambda, \theta, \phi \in \mathbf{R}. \tag{7}$$

Proof. Fix $\theta, \phi \in \mathbf{R}$. For any $\lambda \in \mathbf{R}$, the function $\sigma_\lambda(\,\cdot\,; \theta, \phi)$ is clearly dominated on $[0,1]^n$ by the constant function $g(x) = \max_{t\in[-T,T]} |\sigma(t)|$. The following statements are true.

1. Open subsets of $[0,1]^n$ (in the relative topology of $[0,1]^n$) are μ-measurable by the definition of a Borel measure.
2. g is constant, so it is continuous, meaning that $g^{-1}(U)$ is open for any open set U (in the relative topology of $[0,1]^n$).
3. 1. and 2. imply that $g^{-1}(U)$ is μ-measurable for any open set U (in the relative topology of $[0,1]^n$).
4. 3. implies that g is μ-measurable.
5. $[0,1]^n$ is open in the relative topology on $[0,1]^n$, so it is μ-measurable by 1.
6. g is constant on the μ-measurable set $[0,1]^n$, so g is simple with respect to the measure μ by definition.
7. Simple functions are always integrable, so, by 6., g is integrable with respect to μ.
8. σ is continuous, and $x \mapsto \lambda(y^T x + \theta) + \phi$ is also continuous, so $\sigma_\lambda(\,\cdot\,; \theta, \phi)$, the composition, of the two, is also continuous.
9. By 8., $\sigma_\lambda(\,\cdot\,; \theta, \phi)$ is continuous on $[0,1]^n$, so $\sigma_\lambda^{-1}(U; \theta, \phi)$ is open for any open set U (in the relative topology of $[0,1]^n$).
10. 1. and 9. imply that $\sigma_\lambda^{-1}(U; \theta, \phi)$ is open for any open set U (in the relative topology of $[0,1]^n$).
11. 10. implies that $\sigma_\lambda(\,\cdot\,; \theta, \phi)$ is μ-measurable.

Let $\{\lambda_n\}$ be any sequence of real numbers such that $\lambda_n \to \infty$ as $n \to \infty$. Then $\sigma_{\lambda_n}(x; \theta, \phi) \to \gamma(x)$ as $n \to \infty$ for all $x \in [0,1]^n$ by the previous problem, and $\{\sigma_{\lambda_n}(\,\cdot\,; \theta, \phi)\}$ and g satisfy the hypotheses of the Dominated Convergence Theorem, so

$$\begin{aligned}
0 &= \lim_{n\to\infty} \int_{[0,1]^n} \sigma_{\lambda_n}(x; \theta, \phi)\, \mathrm{d}\mu(x) && (8)\\
&= \int_{[0,1]^n} \gamma(x)\, \mathrm{d}\mu(x) && (9)\\
&= \int_{[0,1]^n \setminus (\Pi_{y,\theta} \cup H_{y,\theta})} \gamma(x)\, \mathrm{d}\mu(x) + \int_{\Pi_{y,\theta}} \gamma(x)\, \mathrm{d}\mu(x) + \int_{H_{y,\theta}} \gamma(x)\, \mathrm{d}\mu(x) && (10)\\
&= \int_{[0,1]^n \setminus (\Pi_{y,\theta} \cup H_{y,\theta})} 0\, \mathrm{d}\mu(x) + \int_{\Pi_{y,\theta}} \sigma(\phi)\, \mathrm{d}\mu(x) + \int_{H_{y,\theta}} 1\, \mathrm{d}\mu(x) && (11)\\
&= \sigma(\phi)\mu(\Pi_{y,\theta}) + \mu(H_{y,\theta}). && (12)
\end{aligned}$$

□

Question 3.

Suppose that μ satisfies (6). Then $\mu = 0$.

Proof. Define the linear functional $F : L^\infty(\mathbf{R}) \to \mathbf{R}$ by

$$F(h) = \int_{[0,1]^n} h\left(y^T x\right)\, \mathrm{d}\mu(x) \tag{13}$$

First, let $h = \chi_{[a,\infty)}$ for some $a \in \mathbf{R}$. Choose $\theta = -a$, and $\phi = T$. Then $\sigma(\phi) = 1$, and

$$h(y^T x) = 1 \iff y^T x \geq a \iff y^T x + \theta \geq 0 \iff y^T x \in \Pi_{y,\theta} \cup H_{y,\theta}. \tag{14}$$

If $y^T x \notin \Pi_{y,\theta} \cup H_{y,\theta}$, then $h(y^T x) = \chi_{[a,\infty)}(y^T x) = 0$, because characteristic functions can only be 0 or 1. Therefore, by the previous problem,

$$\begin{aligned} F(h) &= \int_{[0,1]^n} h(y^T x)\, \mathrm{d}\mu(x) && (15)\\ &= \int_{\Pi_{y,\theta}} 1\, \mathrm{d}\mu(x) + \int_{H_{y,\theta}} 1\, \mathrm{d}\mu(x) && (16)\\ &= 1 \cdot \mu(\Pi_{y,\theta}) + \mu(H_{y,\theta}) = \sigma(\phi)\mu(\Pi_{y,\theta}) + \mu(H_{y,\theta}) = 0. && (17) \end{aligned}$$

Second, for $a, b \in \mathbf{R}$, the characteristic function $\chi_{[a,b)}$ can be written as $\chi_{[a,\infty)} - \chi_{[b,\infty)}$. Since F is linear, it follows that $F\left(\chi_{[a,b)}\right) = 0$. □

Question 4.

Let $\mathcal{N}$ be the set of all functions $G \in C[0,1]$ of the form

$$G(x) = \sum_{j=1}^{n} \alpha_j \sigma(y_j x + \theta_j), \tag{18}$$

for some $\alpha_j, y_j, \theta_j \in \mathbf{R}$. Then $\mathcal{N}$ is dense in $C[0,1]$ in the uniform norm.

Proof. It is clear that $\mathcal{N}$ is a subspace of $C[0,1]$. Then $\overline{\mathcal{N}}$ is a closed subspace of $C[0,1]$.

Suppose that $\mathcal{N}$ is *not* dense in $C[0,1]$. Then there exists some function $f \in C[0,1]$ such that $f \notin \overline{\mathcal{N}}$. Define the linear functional μ_0 on the subspace $\operatorname{span}\{\overline{\mathcal{N}}, f\}$ by setting

$$\mu_0(G + af) = a. \tag{19}$$

This is well-defined because every element $g \in \operatorname{span}\{\overline{\mathcal{N}}, f\}$ can be written uniquely in the form $g = G + af$, where $G \in \mathcal{N}$, and $a \in \mathbf{R}$. Indeed, if we had $G, G' \in \mathcal{N}$, and $a, a' \in \mathbf{R}$ such that $G + af = G' + a'f$, then $G - G' = (a' - a)f$, and $a \neq a'$ would imply that $f \in \mathcal{N}$, so $a = a'$, and $G = G'$.

The functional μ_0 is clearly linear because

$$\mu_0(r(G + af) + s(G' + a'f)) = \mu_0((rG + sG') + (sa + ra')f) = sa + ra' = s\mu_0(G + af) + r\mu_0(G' + a'f). \tag{20}$$

Furthermore, $\mu_0(\mathcal{N}) = \{0\}$ because

$$\mu_0(G) = \mu_0(G + 0f) = 0. \tag{21}$$

Finally, μ_0 is dominated by the sublinear functional $p = \frac{\|\cdot\|}{\|f\|}$ because

$$|\mu_0(G + af)| = |a| = \frac{\|af\|}{\|f\|} \leq \frac{\|G + af\| + \|G\|}{\|f\|} \tag{22}$$

□