Math 6417 Homework 2

Jacob Hauck

October 19, 2023

Question 1.

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.

1.1) Fix $y, \theta, \phi \in \mathbf{R}$, and define

$$\sigma_\lambda(x) = \sigma(\lambda(yx + \theta) + \phi), \qquad x \in \mathbf{R}. \tag{2}$$

Then

$$\lim_{\lambda \to \infty} \sigma_\lambda(x) = \gamma(x) = \begin{cases} 1 & yx + \theta > 0 \\ 0 & yx + \theta < 0 \\ \sigma(\phi) & yx + \theta = 0. \end{cases} \tag{3}$$

Proof. We use proof by cases. For any $x \in \mathbf{R}$, there are exactly three possibilities.

I. $yx + \theta > 0$.
Set $\Lambda = \frac{T - \phi}{yx + \theta}$. Then $\lambda \geq \Lambda$ implies that $\lambda(yx + \theta) + \phi \geq T$, so $\sigma_\lambda(x) = 1$ because σ is sigmoidal. Therefore, $\sigma_\lambda(x) \to 1 = \gamma(x)$ as $\lambda \to \infty$.

II. $yx + \theta < 0$.
Set $\Lambda = \frac{-T - \phi}{yx + \theta}$. Then $\lambda \geq \Lambda$ implies that $\lambda(yx + \theta) + \phi \leq -T$, so $\sigma_\lambda(x) = 0$ because σ is sigmoidal. Therefore, $\sigma_\lambda(x) \to 0 = \gamma(x)$ as $\lambda \to \infty$.

III. $yx + \theta = 0$.
In this case, $\sigma_\lambda(x) = \sigma(\phi)$ for all λ, so clearly $\sigma_\lambda(x) \to \sigma(\phi) = \gamma(x)$ as $\lambda \to \infty$.

□

1.2) For any $y, \theta \in \mathbf{R}$, define

$$\Pi_{y,\theta} = \{x \in [0, 1] : yx + \theta = 0\}, \qquad H_{y,\theta} = \{x \in [0, 1] : yx + \theta > 0\}. \tag{4}$$

If μ is a finite Borel measure on $[0, 1]$, then we say that μ has property $(*)$ if and only if

$$\int_0^1 \sigma_\lambda(x) \, \mathrm{d}\mu(x) = 0 \quad \text{for all } \lambda, y, \theta, \phi \in \mathbf{R}. \tag{$*$}$$

If μ has property $(*)$, then

$$0 = \mu(H_{y,\theta}) + \sigma(\phi)\mu(\Pi_{y,\theta}) \quad \text{for all } y, \theta, \phi \in \mathbf{R}. \tag{5}$$

Proof. Let $y, \theta, \phi \in \mathbf{R}$ be given. Since σ is continuous, it is bounded on $[-T, T]$, say by $M > 0$. Since σ is sigmoidal, it follows that $\{\sigma(t) : t \notin [-T, T]\} = \{0, 1\} \subseteq \{\sigma(t) : t \in [-T, T]\}$, so $|\sigma(t)| \leq M$ for all $t \in \mathbf{R}$. This implies that $|\sigma_\lambda(x)| = |\sigma(\lambda(yx + \theta) + \phi)| \leq M$ for all $x \in [0, 1]$ and all λ.

The constant function M is integrable on $[0, 1]$, and $\sigma_\lambda \to \gamma$ pointwise on $[0, 1]$ as $\lambda \to \infty$ by the previous part, so, by the Dominated Convergence Theorem[1],

$$\lim_{\lambda \to \infty} \int_0^1 \sigma_\lambda \, \mathrm{d}\mu = \int_0^1 \gamma \, \mathrm{d}\mu. \tag{6}$$

Define $N_{y,\theta} = [0, 1] \setminus (H_{y,\theta} \cup \Pi_{y,\theta})$, so that $[0, 1] = H_{y,\theta} \cup \Pi_{y,\theta} \cup N_{y,\theta}$ disjointly. Then[2]

$$\begin{aligned} 0 = \int_0^1 \gamma \, \mathrm{d}\mu &= \int_{H_{y,\theta}} \gamma \, \mathrm{d}\mu + \int_{\Pi_{y,\theta}} \gamma \, \mathrm{d}\mu + \int_{N_{y,\theta}} \gamma \, \mathrm{d}\mu &(7) \\ &= \mu(H_{y,\theta}) + \sigma(\phi)\mu(\Pi_{y,\theta}) &(8) \end{aligned}$$

because $\gamma\big|_{H_{y,\theta}} = 1$, $\gamma\big|_{\Pi_{y,\theta}} = \sigma(\phi)$, and $\gamma\big|_{N_{y,\theta}} = 0$ by definition. □

1.3) If μ has property $(*)$, then $\mu = 0$.

Proof. Let μ have property $(*)$. If we choose $\phi = -T$, then $\sigma(\phi) = 0$, so by the previous part, $\mu(H_{y,\theta}) = 0$ for all $y, \theta \in \mathbf{R}$. Then, if we choose $\phi = T$ so that $\sigma(\phi) = 1$, we see that, by the previous part, $\mu(\Pi_{y,\theta}) = 0$ for all $y, \theta \in \mathbf{R}$, as well.

A simple series of deductions shows that[3] $\mu([a, b)) = 0$ for all $0 \leq a \leq b \leq 1$:

- If we choose $y = 0$ and $\theta = 0$, then $\Pi_{y,\theta} = [0, 1]$, so $\mu([0, 1]) = 0$.
- Given $a \in [0, 1]$, if we choose $y = -1$ and $\theta = a$, then $H_{y,\theta} = [0, a)$, so $\mu([0, a)) = 0$.
- Hence, given $a \in [0, 1]$, we have $0 = \mu([0, 1]) = \mu([0, a)) + \mu([a, 1]) = 0 + \mu([a, 1])$, so $\mu([a, 1]) = 0$.
- Finally, given $0 \leq a \leq b \leq 1$, we have $0 = \mu([a, 1]) = \mu([a, b)) + \mu([b, 1]) = \mu([a, b)) + 0$, so $\mu([a, b)) = 0$.

The set of finite Borel measures on $[0, 1]$ can be shown to be isomorphic as a Banach space[4] to $(C([0, 1]))^*$ by defining the action of a measure μ on a continuous function $h \in C([0, 1])$ by

$$\mu(h) = \int_0^1 h \, \mathrm{d}\mu. \tag{9}$$

Hence, we only need to show that $\mu = 0$ as an element of $(C([0, 1]))^*$, that is, $\mu(h) = 0$ for all $h \in C([0, 1])$.

We start by showing that $\mu\left(\chi_{[a,b)}\right) = 0$ for all $0 \leq a \leq b \leq 1$. This is easy enough because

$$\mu\left(\chi_{[a,b)}\right) = \int_0^1 \chi_{[a,b)} \, \mathrm{d}\mu = \mu([a, b)) = 0. \tag{10}$$

Given $h \in C([0, 1])$ and a natural number n, we can choose $\delta > 0$ such that $|h(t) - h(s)| < \frac{1}{n}$ for all $t, s \in [0, 1]$ satisfying $|t - s| < \delta$. Choose a partition $\{t_i\}_{i=1}^{N_n}$ of $[0, 1]$ such that $t_i - t_{i-1} < \delta$ for $i = 1, 2, \ldots, N_n$.

[1] σ_λ is inetgrable because it is continuous
[2] Each of $H_{y,\theta}$, $\Pi_{y,\theta}$ and $N_{y,\theta}$ is measurable because they are just intervals, and μ is Borel
[3] $[a, b)$ is measurable because μ is Borel
[4] Using the uniform norm in $C([0, 1])$

Define the function

$$h_n(t) = \begin{cases} \sum_{i=1}^{N_n} m_i \cdot \chi_{[t_{i-1},t_i)}(t) & t \in [0,1) \\ h(1) & t = 1, \end{cases} \tag{11}$$

where $m_i = \min_{t \in [t_{i-1},t_i]} h(t)$. Since h is continuous, it achieves the value m_i for some $t^* \in [t_{i-1}, t_i]$. On the other hand, $h_n(t) = m_i$ for all $t \in [t_{i-1}, t_i)$. For all $t \in [t_{i-1}, t_i)$, we have $|t^* - t| < \delta$ by the construction of the partition, so for all $t \in [t_{i-1}, t_i)$,

$$|h(t) - h_n(t)| = |h(t) - m_i| = |h(t) - h(t^*)| < \frac{1}{n}. \tag{12}$$

Since this holds for all i, and $h_n(1) = h(1)$ for all n, it follows that $h_n \to h$ uniformly on $[0, 1]$.

Clearly h_n is simple by construction, so h_n is measurable for all n. Therefore, by the uniform convergence of h_n to h,

$$\begin{aligned} \mu(h) &= \int_0^1 h \, \mathrm{d}\mu = \lim_{n\to\infty} \int_0^1 h_n \, \mathrm{d}\mu && (13) \\ &= \lim_{n\to\infty} \left[h(1) \int_{\{1\}} \mathrm{d}\mu + \sum_{i=1}^{N_n} m_i \int_{[t_{i-1},t_i)} \chi_{[t_{i-1},t_i)} \, \mathrm{d}\mu \right] && (14) \\ &= \lim_{n\to\infty} \left[h(1)\mu(\{1\}) + \sum_{i=1}^{N_n} m_i \mu([t_{i-1}, t_i)) \right] && (15) \\ &= \lim_{n\to\infty} h(1)\mu([1,1]) && (16) \\ &= \lim_{n\to\infty} 0 = 0 && (17) \end{aligned}$$

because $\mu([t_{i-1}, t_i)) = 0$ for all i, and $\mu(\{1\}) = \mu([1, 1]) = 0$ by the third bullet near the beginning of the proof.

Thus, $\mu(h) = 0$ for all $h \in C([0, 1])$, that is, $\mu = 0$. □

1.4) Let $\mathcal{N}$ denote the set of functions $G : [0, 1] \to \mathbf{R}$ of the form

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

for some positive integer N and parameters $\alpha_j, y_j, \theta_j \in \mathbf{R}$, $j = 1, 2, \ldots, N$. Then $\mathcal{N}$ is dense in $C([0, 1])$ in the uniform norm.

Proof. Since σ is continuous, it follows easily that G is continuous for all $G \in \mathcal{N}$. Hence, $\mathcal{N} \subseteq C([0, 1])$. It is obvious that $\mathcal{N}$ is nonempty, but $\mathcal{N}$ is also closed under linear combinations of elements of $\mathcal{N}$ because, given $G, G' \in \mathcal{N}$ and $p, p' \in \mathbf{R}$, there exist positive integers N, N', parameters α_j, y_j, θ_j, $j = 1, 2, \ldots N$, and parameters $\alpha'_j, y'_j, \theta'_j$, $j = 1, 2, \ldots, N'$ such that

$$\begin{aligned} pG + p'G' &= p \sum_{j=1}^{N} \alpha_j \sigma(y_j x + \theta_j) + p' \sum_{j=1}^{N'} \alpha'_j \sigma(y'_j x + \theta'_j) && (19) \\ &= \sum_{j=1}^{N+N'} \beta_j \sigma(z_j x + \vartheta_j), && (20) \end{aligned}$$

taking $\beta_j = p\alpha_j$, $z_j = y_j$, $\vartheta_j = \theta_j$ if $j \leq N$, and $\beta_j = p'\alpha'_{j-N}$, $z_j = y'_{j-N}$, $\vartheta_j = \theta'_{j-N}$ if $j > N$. Thus, $pG + p'G' \in \mathcal{N}$ by definition, and $\mathcal{N}$ is a vector subspace of $C([0,1])$.

Then $\overline{\mathcal{N}}$ is also a subspace. Suppose that $\overline{\mathcal{N}} \neq C([0,1])$. Then we can choose some $g \in C([0,1]) \setminus \overline{\mathcal{N}}$. Let $d = \mathrm{dist}(g, \overline{\mathcal{N}})$. Then $d > 0$ because $\overline{\mathcal{N}}$ is closed. By the Mazur separation lemma 1, there exists some $\mu \in (C([0,1]))^*$ such that $\mu(g) = d$, and $\mu(G) = 0$ for all $G \in \overline{\mathcal{N}}$.

Clearly $\sigma_\lambda \in \mathcal{N} \subseteq \overline{\mathcal{N}}$ for any choice of $\lambda, y, \theta, \phi \in \mathbf{R}$; thus, $\mu(\sigma_\lambda) = 0$ for all σ_λ. Using the fact that μ is a finite Borel measure on $[0,1]$ (as mentioned in the previous part) with

$$\mu(h) = \int_0^1 h \, \mathrm{d}\mu \quad \text{for all } h \in C([0,1]), \tag{21}$$

we see that

$$0 = \mu(\sigma_\lambda) = \int_0^1 \sigma_\lambda \, \mathrm{d}\mu \tag{22}$$

for all σ_λ. That is, μ has property $(*)$. Therefore, by the previous part, $\mu = 0$. This contradicts the fact that $\mu(g) = d \neq 0$. Hence, $\overline{\mathcal{N}} = C([0,1])$, which is equivalent to saying that $\mathcal{N}$ is dense in $C([0,1])$. □

Question 2.

For any positive integer n, define the linear functional $\ell_n : C([0,1]) \to \mathbf{R}$ by

$$\ell_n(f) = \sum_{j=0}^{n} w_j^n f(x_j^n), \tag{23}$$

where $\{x_j^n\}_{j=0}^n$ is a partition of $[0,1]$, and $\{w_j^n\}_{j=0}^n$ is a sequence of real numbers. $\ell_n(f)$ is meant to be a numerical quadrature formula for the integral of f on $[0,1]$ with weight function $w \in L^1([0,1])$.

2.1) Equipping $C([0,1])$ with the uniform norm, the functional ℓ_n is a bounded, linear functional, and the induced norm of ℓ_n is given by

$$\|\ell_n\| = \sum_{j=0}^{n} |w_j^n|. \tag{24}$$

Proof. Let $f, g \in C([0,1])$, and let $\alpha, \beta \in \mathbf{R}$. Then

$$\begin{aligned}\ell_n(\alpha f + \beta g) &= \sum_{j=0}^{n} w_j^n(\alpha f(x_j^n) + \beta g(x_j^n)) = \alpha \sum_{j=0}^{n} w_j^n f(x_j^n) + \beta \sum_{j=0}^{n} w_j^n g(x_j^n) &(25)\\ &= \alpha \ell_n(f) + \beta \ell_n(g), &(26)\end{aligned}$$

so ℓ_n is linear.

Let $f \in C([0,1])$ have $\|f\| \leq 1$. Then $|f(x)| \leq 1$ for all $x \in [0,1]$ by the definition of $\|f\|$, and

$$|\ell_n(f)| \leq \sum_{j=0}^{n} |w_j^n| \cdot |f(x_j^n)| \leq \sum_{j=0}^{n} |w_j^n|. \tag{27}$$

Since f was arbitrary with norm bounded by 1,

$$\|\ell_n\| \leq \sum_{j=0}^{n} |w_j^n|. \tag{28}$$

On the other hand, we can choose $f \in C([0,1])$ such that $\|f\| = 1$, and $f(x_j^n) = \operatorname{sgn}(w_j^n)$, the sign of w_j^n, by considering the piecewise linear function that interpolates between the points $\{(x_j^n, \operatorname{sgn}(w_j^n))\}_{j=0}^n$ (if all the w_j^n are zero for some fixed n, then choose $f = 1$). Hence,

$$\|\ell_n\| \geq \frac{|\ell_n(f)|}{\|f\|} = \left|\sum_{j=0}^n w_j^n f(x_j^n)\right| = \sum_{j=0}^n |w_j^n| \tag{29}$$

because $|w_j^n| = w_j^n \cdot \operatorname{sgn}(w_j^n) = w_j^n f(x_j^n)$. Therefore,

$$\|\ell_n\| = \sum_{j=0}^n |w_j^n|. \tag{30}$$

□

2.2) Suppose that the formula converges in the sense that

$$\lim_{n\to\infty} \left| \int_0^1 f(x)w(x)\ \mathrm{d}x - \ell_n(f) \right| = 0 \quad \text{for all } f \in C([0,1]). \tag{31}$$

Then

$$\sup_{n\geq 0} \left(\sum_{j=0}^n |w_j^n| \right) < \infty. \tag{32}$$

Proof. By the Banach-Steinhaus theorem (with $X = C([0,1])$, $Y = \mathbf{R}$, and $T_\alpha = \ell_\alpha$, $\alpha = 1, 2, \ldots$, in the notation of the slides), either $\{\|\ell_n\|\}_{n=1}^\infty$ is bounded, or $\sup\limits_{n\geq 0} |\ell_n(f)| = \infty$ for some $f \in C([0,1])$. The second possibility is false, because, by the convergence assumption (31), the sequence $\{\ell_n(f)\}_{n=1}^\infty$ converges to a finite number for all $f \in C([0,1])$, and convergent sequences are bounded. Therefore, $\{\|\ell_n\|\}$ is bounded.
Since

$$\|\ell_n\| = \sum_{j=0}^n |w_j^n|, \tag{33}$$

it follows that

$$\left\{ \sum_{j=0}^n |w_j^n| \right\}_{n=1}^\infty \tag{34}$$

is bounded, that is,

$$\sup_{n\geq 0} \left(\sum_{j=0}^n |w_j^n| \right) < \infty. \tag{35}$$

□

2.3) Suppose that

$$\sup_{n\geq 0} \left(\sum_{j=0}^n |w_j^n| \right) < \infty. \tag{36}$$

Suppose furthermore that for any polynomial $p(x)$ defined on $[0,1]$

$$\lim_{n\to\infty} \left| \int_0^1 p(x)w(x)\ \mathrm{d}x - \ell_n(p) \right| = 0. \tag{37}$$

Then the quadrature formula ℓ_n works on all of $C([0,1])$ in the sense that

$$\lim_{n\to\infty}\left|\int_0^1 f(x)w(x)\,\mathrm{d}x - \ell_n(f)\right| = 0 \quad \text{for all } f\in C([0,1]). \tag{38}$$

Proof. Let $f\in C([0,1])$, and let $\varepsilon>0$ be given. By the Weierstrass Approximation Theorem, we can choose a polynomial $p(x)$ defined on $[0,1]$ such that

$$\|p-f\|_{C([0,1])} < \varepsilon, \tag{39}$$

where $\|\cdot\|_{C([0,1])}$ is the usual uniform norm.
Furthermore, by (37), we can choose N large enough that

$$\left|\int_0^1 p(x)w(x)\,\mathrm{d}x - \ell_n(p)\right| < \varepsilon \tag{40}$$

for all $n>N$, and, by (36), we can choose $M>0$ such that

$$\sum_{j=0}^n |w_j^n| < M \tag{41}$$

for all $n\geq 0$. Therefore, if $n>N$, then

$$\begin{aligned}
\left|\int_0^1 f(x)w(x)\,\mathrm{d}x - \ell_n(f)\right| &\leq \left|\int_0^1 (f(x)-p(x))w(x)\,\mathrm{d}x\right| && (42)\\
&\quad + |\ell_n(p)-\ell_n(f)| + \left|\int_0^1 p(x)w(x)\,\mathrm{d}x - \ell_n(p)\right| && (43)\\
&\leq \varepsilon\int_0^1 |w(x)|\,\mathrm{d}x + \sum_{j=0}^n |p(x_j^n)-f(x_j^n)|\cdot|w_j^n| + \varepsilon && (44)\\
&\leq (\|w\|_{L^1}+M+1)\varepsilon. && (45)
\end{aligned}$$

Since $\|w\|_{L^1}+M+1$ is independent of n and $\varepsilon>0$ was arbitrary, it follows that

$$\lim_{n\to\infty}\left|\int_0^1 f(x)w(x)\,\mathrm{d}x - \ell_n(f)\right| = 0. \tag{46}$$

Thus, the quadrature rule works for the arbitrary continuous function $f\in C([0,1])$, and, consequently, for all functions in $C([0,1])$. $\square$