Topics in Applied Mathematics: Infinitely Large Neural Networks, 3341.751
E. Ryu
Spring 2022

Homework 7
Due 5pm, Friday, June 17, 2022

Problem 1: *Solution via test functions.* Let $f\colon (0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ be measurable. Consider the ODE

$$\dot{X} = f(t, X), \qquad X(0) = X_0.$$

We define $X\colon [0,\infty) \to \mathbb{R}$ to be a solution if X is measurable and

$$X(t) = X_0 + \int_0^t f(s, X(s))\, ds, \qquad \text{for } t \geq 0.$$

Show that X is a solution if and only if

$$\int_0^\infty \varphi(t) f(t, X(t))\, dt = -\int_0^\infty \varphi'(t) X(t)\, dt, \qquad \forall\, \varphi \in \mathcal{C}_c^\infty((0,\infty))$$

Clarification. $\mathcal{C}_c^\infty(0,\infty)$ is the set of infinitely smooth functions with compact support in $(0,\infty)$. Thus, $\lim_{t\to 0_+} \varphi(t) = 0$ for all $\varphi \in \mathcal{C}_c^\infty((0,\infty))$. (So $\mathcal{C}_c^\infty((0,\infty))$ and $\mathcal{C}_c^\infty([0,\infty))$ are materially different.)

Problem 2: *Discontinuous weak solutions of a PDE.* Consider the PDE

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \qquad u(0,\cdot) = \mathbf{1}_{[-1,1]}.$$

Show that

$$u(t,x) = \mathbf{1}_{[-1,1]}(x-t)$$

is a weak solution, i.e., show that

$$0 = \int_{-\infty}^\infty \int_0^\infty (\partial_t \varphi(t,x) + \partial_x \varphi(t,x)) u(t,x)\, dtdx, \quad \forall\, \varphi \in C_c^\infty((0,\infty)\times\mathbb{R}).$$

Remark. Even though $u(t,x)$ is decided not differentiable, it is a (weak) solution of a partial *differential* equation.

Problem 3: *Divergence theorem for the continuity equation.* Assume $\rho(t,x)$ is continuously differentiable in t and x. Assume $\mathbf{v}(t,x;\rho(t,\cdot))$ is continuously differentiable in x. Let $\varphi \in C_c^\infty((0,\infty)\times\mathbb{R}^d)$. Show that

$$\int_{\mathbb{R}^d}\int_0^\infty \varphi \partial_t \rho\, dtdx + \int_0^\infty \int_{\mathbb{R}^d} \varphi \nabla\cdot(\rho \mathbf{v})\, dxdt$$
$$= -\int_{\mathbb{R}^d}\int_0^\infty \partial_t\varphi\rho\, dtdx - \int_0^\infty\int_{\mathbb{R}^d} \rho((\nabla\varphi)\cdot\mathbf{v})\, dxdt.$$

Problem 4: *Exercise on functional derivatives I.* Let $\Phi\colon \mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\to\mathbb{R}$ be a bounded measurable function. (Boundedness is assumed to ensure all integrals are well defined.) Let $P\in\mathcal{P}(\mathbb{R}^{d_1})$ and define

$$U(\theta,\theta') = \mathbb{E}_{X\sim P}\left[\Phi(X;\theta)\Phi(X;\theta')\right]$$
$$V(\theta) = \mathbb{E}_{X\sim P}\left[\Phi(X;\theta)f_\star(X)\right].$$

Define $\tilde{R}\colon \mathcal{P}(\mathbb{R}^{d+2})\to\mathbb{R}$ as

$$\tilde{R}(\rho)=\frac{1}{2}\int_{\mathbb{R}^{d_2}}\int_{\mathbb{R}^{d_2}} U(\theta,\theta')\,d\rho(\theta)d\rho(\theta') - \int_{\mathbb{R}^{d_2}} V(\theta)\,d\rho(\theta).$$

Show that

$$\frac{\delta\tilde{R}}{\delta\rho}\Big|_\rho(\cdot) = \int U(\cdot,\theta')\,d\rho(\theta') - V(\cdot).$$

Problem 5: *Exercise on functional derivatives II.* Let $R\colon\mathcal{H}\to\mathbb{R}$ be Fréchet differentiable, and write $\partial_f R|_{f_0}\in\mathcal{H}$ for the Fréchet derivative of R at f_0. Let $\Phi(\cdot;\theta)\in\mathcal{H}$ for all $\theta\in\Theta$. (The primary example we consider is $\Phi(x;\theta)=u\sigma(a^\intercal x+b)$, where σ is a nonlinear activation function and $\theta=(u,a,b)$.) Define $\tilde{R}\colon\mathcal{P}(\Theta)\to\mathbb{R}$ as

$$\tilde{R}(\rho)=R\left[\int_\Theta \Phi(\cdot;\theta)\,d\rho(\theta)\right].$$

Show that

$$\frac{\delta\tilde{R}}{\delta\rho}\Big|_\rho(\theta)=\left\langle \partial_f R|_{\int_\Theta \Phi(\cdot;\theta)\,d\rho(\theta)},\Phi(\cdot;\theta)\right\rangle_{\mathcal{H}},$$

which is often expressed more concisely as

$$\frac{\delta\tilde{R}}{\delta\rho}\Big|_\rho(\theta)=\left\langle \partial_f R|_{\int_\Theta \Phi\,d\rho},\Phi(\cdot;\theta)\right\rangle_{\mathcal{H}},$$

or

$$\frac{\delta\tilde{R}}{\delta\rho}=\langle\partial_f R,\Phi\rangle_{\mathcal{H}}.$$

Problem 6: *Continuity equation requires continuous flow.* Let $\eta\in\mathcal{C}_c^\infty(\mathbb{R})$ satisfy $\eta\ge 0$, $\eta(x)=0$ for $|x|\ge 1/3$, and

$$\int_{-1/3}^{1/3}\eta(x)\,dx=1.$$

Let

$$\rho(t,x)=(1-t)\eta(x)+t\eta(x-1),\qquad \text{for } t\in[0,1],\ x\in\mathbb{R}.$$

Show that ρ cannot be a solution of the continuity equation

$$\partial_t\rho+\nabla\cdot(\rho\mathbf{v})=0$$

no matter the choice of $\mathbf{v}$.

Hint. Integrate with respect to $x\in[-1/2,1/2]$ and use the divergence theorem.

Remark. The point is that the continuity equation only allows continuous flow. Particles or mass are not allowed to “teleport”.

Problem 7: *Non-continuous transport has infinite curve length under W_2.* Consider $\gamma_1\colon [0,1] \to \mathcal{P}(\mathbb{R})$ defined as

$$\gamma_1(t) = \delta_t$$

and $\gamma_2\colon [0,1] \to \mathcal{P}(\mathbb{R})$ defined as

$$\gamma_2(t) = (1-t)\delta_0 + t\delta_1.$$

Under the W_2 metric, show that γ_1 has length $1/\sqrt{2}$ while γ_2 has length ∞.

Remark. In fact, γ_1 is a constant-speed geodesic.

Remark. For a metric space (X,d) and a continuous function $\gamma\colon [0,1] \to X$, the length of the curve γ is defined as

$$\sup_{\substack{0=t_0<t_1<\cdots<t_n=1 \\ n\in\mathbb{N}}} \sum_{i=0}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})).$$

Problem 8: *Divergence theorem on weak solutions.* Let $\{\rho_t\}_{t\geq 0} \subset \mathcal{P}(\mathbb{R}^{d+2})$ be a weak solution to

$$\partial_t \rho_t = \operatorname{div}(\rho_t \nabla J(\cdot|\rho_t)).$$

Assume that $\{\rho_t\}_{t\geq 0}$ is (weakly) continuous in t, i.e., for all $t_0 \in [0,\infty)$ and $h \in \mathcal{C}_c^\infty(\mathbb{R}^{d+2})$,

$$\lim_{t\to t_0} \int_{\mathbb{R}^{d+2}} h(\theta)\, d\rho_t(\theta) = \int_{\mathbb{R}^{d+2}} h(\theta)\, d\rho_{t_0}(\theta).$$

Show that for all $0 \leq t_1 < t_2$ and $h \in \mathcal{C}_c^\infty(\mathbb{R}^{d+2})$,

$$\int_{\mathbb{R}^{d+2}} h(\theta)\, d(\rho_{t_2} - \rho_{t_1})(\theta) = -\int_{t_1}^{t_2} \int_{\mathbb{R}^{d+2}} \nabla h(\theta) \cdot \nabla J(\theta|\rho_t)\, d\rho_t(\theta) dt.$$

Remark. One can prove that $\{\rho_t\}_{t\geq 0}$ is (or can be chosen to be) continuous, instead of assuming the property.

Remark. With formal calculations, we can "derive"

$$\rho_{t_2} - \rho_{t_1} = \int_{t_1}^{t_2} \partial_t \rho_t\, dt = \int_{t_1}^{t_2} \operatorname{div}(\rho_t \nabla J(\cdot|\rho_t))\, dt.$$

However, this is not rigorous since the divergence theorem (or the fundamental theorem of calculus) $\rho_{t_2} - \rho_{t_1} = \int_{t_1}^{t_2} \partial_t \rho_t\, dt$ has not been established.

Hint. Let

$$\Psi(t) = \begin{cases} \exp\left(-\frac{1}{1-t^2}\right) & \text{for } t \in (-1,1) \\ 0 & \text{otherwise,} \end{cases} \qquad \phi_\delta(t) = \frac{1}{\delta \int_{\mathbb{R}} \Psi(t)\, dt} \Psi(t/\delta).$$

(Then $\phi_\delta \in \mathcal{C}^\infty$ is supported on $[-\delta,\delta]$ and $\int_{\mathbb{R}} \phi_\delta(t)\, dt = 1$.) Let

$$g_\delta(t) = \int_{-\infty}^{t} \phi_\delta(t-t_1) - \phi_\delta(t-t_2)\, dt.$$

Let

$$\varphi_\delta(t,\theta) = g_\delta(t) h(\theta) \in \mathcal{C}_c^\infty((0,\infty) \times \mathbb{R}^{d+2}),$$

and consider $\delta \to 0$.