CAAM/STAT 31310, Autumn 2024, U Chicago

CAAM 31310: Homework 1

Due Oct 15, '24 (11:59 pm ET) on Gradescope

Cite any sources and collaborators; do not copy. See syllabus for policy.

Problem 1

Take X_i to be independent zero-mean random variables with unit variance. Usually, the classical central limit theorem (CLT) is proved by showing that the characteristic function of $Y_n := (1/\sqrt{n})\sum_{i=1}^n X_i$ converges to the characteristic function of a standard normal. Our task in this problem is to use Stein's identity to prove the CLT. Stein's method has been used to derive computational methods for bounding the distance between two random variables and for a class of sampling algorithms, where the task is to generate samples from a partially specified probability density.

(1) First show that if W is a standard normal RV, for any bounded, differentiable function $f : \mathbb{R} \to \mathbb{R}$, with $\mathbb{E}[Wf(W)], \mathbb{E}f'(W) \leq \infty$, $\mathcal{A}f(W) := f'(W) - Wf(W)$ has zero mean (1 point). This is known as *Stein's lemma*.

(2) For any function $g \in \mathrm{Lip}_1(\mathbb{R})$ (differentiable functions with Lipschitz constant = 1), there is a solution f to *Stein's equation*: $\mathcal{A}f(x) = g(x) - \mathbb{E}g(W)$, where W is a standard normal. (2 points)

(3) Show that a bounded solution f exists that is twice-differentiable with $\|f''\|_\infty \leq 2$ and $\|f'\|_\infty \leq \sqrt{\pi}/2$. (2 points)

(4) Let $\mathbb{E}|X_i|^3 \leq \infty$. Using (2) and (3) above, find a function class $\mathcal{F}$ such that the Wasserstein-1 distance,

$$W^1(Y_n, W) := \sup_{g \in \mathrm{Lip}_1(\mathbb{R})} |\mathbb{E}g(W) - \mathbb{E}g(Y)| \leq \sup_{f \in \mathcal{F}} |\mathbb{E}\mathcal{A}f(Y_n)| \leq C\mathbb{E}|X_i^3|/\sqrt{n}.$$

(2 points)

(6) Use (5) and the converse of Stein's lemma to prove the CLT: $Y_n \xrightarrow{\mathrm{d}} W$. (2 points)

Problem 2

This problem asks you to think about an iterative numerical method as a discrete-time dynamical system (map). Consider the power iteration method for a square, non-singular, diagonalizable matrix $A \in \mathbb{R}^{d\times d}$. For $t \in \mathbb{N}$,

- $v_t \to Av_{t-1}$

(*) $v_{t+1} \to v_{t+1}/\|v_{t+1}\|$.

1. Write down a map $F(x_t) = x_{t+1}$ to describe the above algorithm, where F is defined on a set $M \subseteq \mathbb{R}^d$. (1 point)
2. Is M compact? (1 point)
3. Is F a contraction on M? (1 point)
4. How many fixed points does F have? (1 point) What are they? (1 point)
5. State the assumptions on A so that almost every initial condition converges to a fixed point. (1 point)
6. Under the assumptions in the part above, prove the convergence of almost every iterate to a fixed point of F. (3 points)
7. From here on, consider the power iteration without the normalization step (*). Write the corresponding new map, F, on $\mathbb{R}^d$ (1 point).
8. Give conditions on A for F to be a contraction map (1 point).
9. Give conditions on A for F to be a linear hyperbolic map (1 point).
10. Without the additional conditions in the above two parts (i.e., without hyperbolicity assumptions), describe the asymptotic behavior of all orbits of F. That is, give, with justification, a stable-unstable-center decomposition of $\mathbb{R}^d$ by F. (3 points).