SINGULAR CLT: PROOF TECHNIQUES

EZRA MILLER
JOINT WITH JONATHAN MATTINGLY AND DO TRAN

INTERACTIONS OF STATISTICS AND GEOMETRY (ISAG) II

NATIONAL UNIVERSITY OF SINGAPORE
14–16 OCTOBER 2024

1. Escape vector characterizations

Theorem 1.1. *Fix an amenable measure μ on a smoothly stratified metric space $\mathcal{M}$, a sequence of measures Δ_n sampled from $T_{\bar{\mu}}\mathcal{M}$ converging weakly to $\Delta \Subset T_{\bar{\mu}}\mathcal{M}$, and a sequence of positive real numbers $t_n \to 0$. The escape vector $\mathscr{E} = \mathscr{E}(\Delta)$ is well defined and confined to the escape cone E_μ in the sense that*

$$\begin{aligned}\mathscr{E}(\Delta) &= \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{X\in E_{\bar{\mu}}^e}\big(F(\exp_{\bar{\mu}} X) + F_{t\delta}(\exp_{\bar{\mu}} X)\big)\\ &= \lim_{t\to 0}\frac{1}{t}\operatorname*{argmin}_{X\in E_{\bar{\mu}}^e}\big(F(\exp_{\bar{\mu}} X) - t\langle\Delta, X\rangle\big)\\ &= \lim_{n\to\infty}\frac{1}{t_n}\operatorname*{argmin}_{X\in E_{\bar{\mu}}^e}\big(F(\exp_{\bar{\mu}} X) - t_n\langle\Delta_n, X\rangle\big)\end{aligned}$$

Remark 1.2. The top line in the display from Theorem 1.1 comes directly from the definition of escape vector, which thinks of escape as minimizing a Fréchet function. Note that the Fréchet function $F_{t\delta}$ is essentially the square distance to δ, so $\nabla_\mu F_{t\delta}(X) = -\langle\Delta, X\rangle$ by [MMT23b, Proposition 2.5]. The transition from the top line in the display to the middle line therefore replaces $F_{t\delta}(\exp_{\bar{\mu}} X)$ with its linear approximation $-t\langle\Delta, X\rangle$. The transition to the bottom line in the display replaces $-t\langle\Delta, X\rangle$, which has a continuously decreasing parameter t and a fixed discrete measure Δ, with a version $t_n\langle\Delta_n, X\rangle$ that has a discrete sequence of parameters decreasing to 0 and varying discrete measures Δ_n converging to Δ.

Remark 1.3. The term $t_n\langle\Delta_n, X\rangle$ morally wants to be the average random tangent field $\overline{g}_n(X)$. The goal is indeed to make that substitution, and also then substitute the Gaussian random tangent field $G(X)$ for $-t\langle\Delta, X\rangle$.

Date: 14–16 October 2024.

Remark 1.4. Theorem 1.1 implies continuity of the escape map $\mathscr{E}$ as a function on measures sampled from $T_{\bar\mu}\mathcal{M}$; see [MMT23d, Corollary 4.32]. More importantly, Theorem 1.1 asserts a family-of-functions continuity: if positive real numbers $t_n \to 0$ are given and

$$\mathscr{E}_n(\Delta) = \frac{1}{t_n} \operatorname*{argmin}_{X \in E^e_{\bar\mu}} \bigl(F(\exp_{\bar\mu} X) - t_n\langle\Delta, X\rangle\bigr)$$

is the n^{th} *escape approximation* of any measure Δ sampled from $T_{\bar\mu}\mathcal{M}$, then Theorem 1.1 asserts that $\mathscr{E}_n(\Delta_n) \to \mathscr{E}(\Delta)$ when $\Delta_n \to \Delta$. A functional version of this convergence is one of the main goals of this lecture, tailored to fit the hypotheses of a particular form of the continuous mapping theorem [VW13, Theorem 1.11.1], for application in the proof of Theorem 2.1.

2. Perturbative CLT

Theorem 2.1 (Perturbative CLT)**.** *Fix a localized immured amenable measure μ on a smoothly stratified metric space $\mathcal{M}$. The empirical Fréchet mean $\bar\mu_n$, empirical and Gaussian tangent perturbations H_n and $H(t)$, and escape vector $\mathscr{E}(\Gamma_\mu)$ of any Gaussian mass Γ_μ satisfy*

$$\begin{aligned}
\lim_{n\to\infty} \sqrt{n} \log_{\bar\mu} \bar\mu_n &\overset{d}{=} \lim_{n\to\infty} \sqrt{n} \operatorname*{argmin}_{X \in T^e_{\bar\mu}\mathcal{M}} \bigl(F(\exp_{\bar\mu} X) - \overline{g}_n(X)\bigr)\\
&\overset{d}{=} \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{X \in \overline{C}^e_{\bar\mu}} \bigl(F(\exp_{\bar\mu} X) - tG(X)\bigr)\\
&= \mathscr{E}(\Gamma_\mu).
\end{aligned}$$

Proof. The top equality is the logarithm of [MMT23d, Proposition 4.24]. It is a direct, hand-dirty calculation via a uniform law of large numbers, which uses amenability for Taylor expansion.

For the bottom equality, by the last line of Theorem 1.1,

$$\begin{aligned}
\mathscr{E}(\Gamma_\mu) &= \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{X \in \overline{C}^e_{\bar\mu}} \bigl(F(\exp_{\bar\mu} X) - t\langle\Gamma_\mu, X\rangle\bigr)\\
&= \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{X \in \overline{C}^e_{\bar\mu}} \bigl(F(\exp_{\bar\mu} X) - tG(X)\bigr).
\end{aligned}$$

where the substitution to reach the second equality is $G(X) = \langle\Gamma_\mu, X\rangle$ (which uses the localized hypothesis). In particular, the limit in the top line exists and is independent of the choice of measurable selection for the Gaussian tangent mass Γ_μ; indeed, already the angular pairing with X does not depend on that selection by [MMT23d, Lemma 6.9]. Consequently, the limit in the bottom line exists and is independent of the path of minimizers represented by the argmin there.

...pause the proof here to introduce representability... □

3. Representable functions and limits

Definition 3.1. Equip the set $\mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ of continuous, positively homogeneous (commuting with nonnegative scaling), real-valued functions on $T_{\bar{\mu}}\mathcal{M}$ with the sup norm

$$\|f\|_\infty = \sup_{Y\in S_{\bar{\mu}}\mathcal{M}} \|f(Y)\|.$$

Functions in $\mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ may be considered as continuous functions on the unit tangent sphere $S_{\bar{\mu}}\mathcal{M}$ without explicit notation to denote restriction to $S_{\bar{\mu}}\mathcal{M}$.

Definition 3.2. A function $R \in \mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ is *representable* if there is a measure Δ sampled from the closure of $\mathbb{R}_+ \operatorname{supp}\widehat{\mu}$ with

$$R(X) = \langle \Delta, X\rangle \text{ for all } X \in E_\mu.$$

A limit $R_n \to R$ in $\mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ is *representable* if all R_n are representable.

Remark 3.3. The term "representable" in Definition 3.2 is meant to evoke the Riesz representation theorem, wherein functionals are represented as inner products.

Example 3.4. The average empirical tangent field $\overline{g}_n$ is representable essentially because $m(\mu, X) = 0$ for $X \in E_\mu$. The Gaussian tangent field G in principle might not be representable, but G is almost surely a limit of representable functions by the CLT for random tangent fields, which is all the extended continuous mapping theorem [VW13, Theorem 1.11.1] needs for application in the proof of Theorem 1.1. Note that even if G is not representable on E_μ as required by Definition 3.2, it is at least representable on the closed fluctuating cone $\overline{C}_\mu$; this major result is the statement from Lecture 4 that $G(X) = \langle \Gamma_\mu, X\rangle$ on the fluctuating cone C_μ.

4. Continuous mapping theorem

Proof of Theorem 1.1, cont'd. For the middle equality, write $\mathscr{E}_0 : \mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R}) \to T_{\bar{\mu}}\mathcal{M}$ for the map

$$\mathscr{E}_0 : R \mapsto \lim_{t\to 0} \frac{1}{t} \operatorname*{argmin}_{X\in \overline{C}{}^e_\mu}\big(F(\exp_{\bar{\mu}} X) - tR(X)\big)$$

defined on the space of functions from Definition 3.1. Denote by $\mathcal{R} \subseteq \mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ the set of representable functions from Definition 3.2 to write $\mathscr{E}_n : \mathcal{R} \to T_{\bar{\mu}}\mathcal{M}$ for the maps

$$\mathscr{E}_n : R \mapsto \frac{1}{t_n} \operatorname*{argmin}_{X\in T^e_{\bar{\mu}}\mathcal{M}}\big(F(\exp_{\bar{\mu}} X) - t_n R(X)\big)$$

for all $n \geq 1$. Another omitted result [MMT23d, Corollary 5.29] says that whenever $R_n \to R$ is a representable limit as in Definition 3.2, $\mathscr{E}_n(R_n) \to \mathscr{E}_0(R)$, where $\mathscr{E}_n$ is defined on $\mathcal{R}$ for $n \geq 1$ and $\mathscr{E}_0$ is defined on $\mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$. Now applying the extended continuous mapping theorem [VW13, Theorem 1.11.1] (an image of this reference is included on the last page of these lecture notes) with the domains $\mathbb{D}_n = \mathcal{R}$ and $\mathbb{D}_0 = \mathcal{C}(T_{\bar{\mu}}\mathcal{M},\mathbb{R})$ to the limit $R_n = \sqrt{n}\,\overline{g}_n \to G = R$ from the random tangent field CLT, which is representable by Example 3.4, yields the desired equality. □

References

[MMT23a] Jonathan Mattingly, Ezra Miller, and Do Tran, *Shadow geometry at singular points of* CAT(κ) *spaces*, preprint, 2023. arXiv:math.MG/2311.09451

[MMT23b] Jonathan Mattingly, Ezra Miller, and Do Tran, *Geometry of measures on smoothly stratified metric spaces*, preprint, 2023. arXiv:math.MG/2311.09453

[MMT23c] Jonathan Mattingly, Ezra Miller, and Do Tran, *A central limit theorem for random tangent fields on stratified spaces*, submitted, 2023. arXiv:math.MG/2311.09454

[MMT23d] Jonathan Mattingly, Ezra Miller, and Do Tran, *Central limit theorems for Fréchet means on stratified spaces*, preprint, 2023. arXiv:math.MG/2311.09455

[VW13] Aad van der Vaart and Jon Wellner, *Weak convergence and empirical processes: with applications to statistics*, Springer Science & Business Media, 2013.

Departments of Mathematics and Statistical Sciences, Duke University, Durham, NC 27708
URL: https://scholars.duke.edu/person/ezra.miller

1.11 Refinements

The continuous mapping theorems for the three modes of stochastic convergence considered so far can be refined to cover maps $g_n(X_n)$, rather than $g(X_n)$, for a fixed g. Then the g_n should have a property that might be called *asymptotic equicontinuity* almost everywhere under the limit measure.

For simplicity, it will be assumed that the limit measure is separable, though this is not necessary for (iii) and can be replaced by other conditions for (i) and (ii) (Problem 1.11.1).

1.11.1 Theorem (Extended continuous mapping). *Let $\mathbb{D}_n \subset \mathbb{D}$ and $g_n: \mathbb{D}_n \mapsto \mathbb{E}$ satisfy the following statements: if $x_n \to x$ with $x_n \in \mathbb{D}_n$ for every n and $x \in \mathbb{D}_0$, then $g_n(x_n) \to g(x)$, where $\mathbb{D}_0 \subset \mathbb{D}$ and $g: \mathbb{D}_0 \mapsto \mathbb{E}$. Let X_n be maps with values in $\mathbb{D}_n$, let X be Borel measurable and separable, and take values in $\mathbb{D}_0$. Then*

(i) $X_n \rightsquigarrow X$ implies that $g_n(X_n) \rightsquigarrow g(X)$;
(ii) $X_n \overset{\mathrm{P}*}{\to} X$ implies that $g_n(X_n) \overset{\mathrm{P}*}{\to} g(X)$;
(iii) $X_n \overset{\mathrm{as}*}{\to} X$ implies that $g_n(X_n) \overset{\mathrm{as}*}{\to} g(X)$.

Proof. Assume the weakest of the three assumptions: the one in (i) that $X_n \rightsquigarrow X$. Let $\mathbb{D}_\infty$ be the set of all x for which there exists a sequence x_n with $x_n \in \mathbb{D}_n$ and $x_n \to x$. First, $\mathrm{P}_*(X \in \mathbb{D}_\infty) = 1$; second, the restriction of g to $\mathbb{D}_0 \cap \mathbb{D}_\infty$ is continuous; and third, if some subsequence satisfies $x_{n'} \to x$ with $x_{n'} \in \mathbb{D}_{n'}$ for every n' and $x \in \mathbb{D}_0 \cap \mathbb{D}_\infty$, then $g_{n'}(x_{n'}) \to g(x)$.

To see the first, invoke the almost sure representation theorem. If $\tilde{X}_n \overset{\mathrm{as}*}{\to} \tilde{X}$ are representing versions, then the range of $\tilde{X}$ is contained in