Notes for Week 7/17/19 - 7/31/19

Alden Green

August 5, 2019

1 Goodness-of-Fit Testing in the Sampling Model.

Let $\mathcal{D} = [0,1]^d$. Suppose we observe the random design $x_1, \ldots, x_n$ independently sampled from the uniform distribution over $[0,1]^d$. Additionally, for $i \in [n]$, we observe

$$z_i = f(x_i) + \varepsilon_i, \quad \varepsilon_i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1), \quad \varepsilon \perp x$$

where $f \in \mathcal{H} \subseteq L^2(\mathcal{D})$. Our statistical goal is hypothesis testing. We wish to distinguish:

$$\mathbf{H}_0 : \|f\|_2 = 0 \quad \text{vs} \quad \mathbf{H}_a : \|f\|_2 > 0.$$

We will evaluate our performance using worst-case error: for a given function class $\mathcal{H}$, test function $\phi : \mathbb{R}^n \to \{0,1\}$ and $\epsilon > 0$, let

$$\mathcal{R}_\epsilon(\phi; \mathcal{H}) = \mathbb{E}_{f=0}(\phi) + \sup_{f \in \mathcal{H} : \|f\|_2 > \epsilon} \mathbb{E}_f(1 - \phi)$$

An example function class $\mathcal{H}$ we will consider will be the unit ball in a Sobolev space. Let s, d be known, fixed positive integers. For $f : \mathcal{D} \to \mathbb{R}$ locally summable, we use the multiindex notation $D^\alpha f$ to denote the αth-weak partial derivative of f (if one exists). Then, the Sobolev norm is

$$\|f\|_{W^{s,2}(\mathcal{D})}^2 = \sum_{|\alpha| \leq s} \int_{\mathcal{D}} \|D^\alpha f\|_2^2 \, dx$$

and the corresponding unit ball is $W^{s,2}(\mathcal{D}; L) = \{f : \|f\|_{W^{s,2}(\mathcal{D})} \leq L\}$.

1.1 Test statistic.

For $x, y \in \mathcal{D}$ and radius $r > 0$ to be specified later, let $\eta_r(x,y) = \mathbf{1}(\|x - y\|/r \leq 1)$. and let A be the $n \times n$ adjacency matrix with entires $A_{ij} = \eta_r(x_i, x_j)$. Let B be the incidence matrix associated with A, and $L = B^T B$ be the associated Laplacian matrix. Write $B = U \Lambda^{1/2} V^T$ for the singular value decomposition of B. Then $L = V \Lambda V^T$ is the eigendecomposition of L , where Λ is a diagonal matrix of eigenvalues with diagonal entries $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, and V is an orthonormal matrix of eigenvectors.

Our test statistic will be the norm of a projection of z onto the subspace spanned by the first few eigenvectors of V. In particular, for $C > 0$ to be specified later, our test statistic will be

$$T_C := \sum_{k : \lambda_k^s \leq C^2} y_k^2, \quad y_k = \frac{1}{\sqrt{n}} z^T v_k.$$

For $b \geq 1$, let $\tau(b) = b\sqrt{2N(C)/n^2}$. Then our test will be $\phi_C = \mathbf{1}\{T_C \leq N(C)/n + \tau(b)\}$.

1.2 Fixed design testing.

Let β be the length-n random vector with jth entry $\beta_j = f(x_j)$. We introduce the scaling $C_{n,r} > 0$, defined by $C_{n,r}^2 = n^2 r^{d+2s}$. We will apply results about testing over balls in discrete Sobolev classes. Let $b \geq 1$ denote $N(C) := \sharp \left\{k : \lambda_k^s \leq C^2\right\}$, and consider the following events:

(E1) Discrete Sobolev norm of β:

$$\|B^{(s)}\beta\|_2 \leq C_{n,r}$$

(E2) Eigenvalue tail decay: For the choice $C = C_r^\star$, where

$$C_r^\star = \frac{(C_{n,r} n^{s/d})^{4s/(4s+d)}}{n^{s/d}},$$

the following inequality is satisfied:

$$N(C_r^\star) \leq n(C_r^\star)^{d/s}.$$

(E3) L_2 norm of β:

$$\frac{\|\beta\|_2^2}{n} \geq \frac{(2\sqrt{2}b + 1)}{n}\left(C_{n,r} n^{s/d}\right)^{2d/(4s+d)}.$$

Suppose for a given $x = x_1, \ldots, x_n$, and for a particular choice of radius r, the events (E3)-(E2) hold. The worst-case error conditional on x can then be upper bounded.

Lemma 1. *For any x and r such that (E1) and (E2) hold, we have that if $f = 0$,*

$$\mathbb{E}_\beta(\phi|x) \leq \frac{1}{b^2}$$

If $f \neq 0$ is such that (E3) is additionally satisfied, we have that

$$\mathbb{E}_\beta(1 - \phi|x) \leq \frac{1}{2b^2} + o(1)$$

We now turn to showing that each of (E1) and (E2) (and, under appropriate conditions on $\|f\|_2$, (E3)) hold with probability at least $1 - \frac{1}{b^2} - o(1)$.

2 Bounding discrete Sobolev norm.

Recall that our goal is to show (E1) occurs with high probability. We will build slowly to this goal.

2.1 $s = 1$, f Lipschitz.

To start, we provide a bound in the case when $s = 1$, and f has bounded Lipschitz norm. Precisely, we define the space of 1-Holder (Lipschitz) functions $C_{\mathcal{X}}^{0,1}(1)$ to consist of all continuous functions $g : \mathcal{D} \to \mathbb{R}$ such that

$$\|g\|_{C_{\mathcal{D}}^{0,1}} := \|g\|_{C(\mathcal{D})} + [g]_{C_{\mathcal{D}}^{0,1}} \leq 1$$

where $\|g\|_{C(\mathcal{D})} = \sup_{x \in \mathcal{D}} |g(x)|$, and

$$[g]_{C_{\mathcal{D}}^{0,1}} = \sup_{x,y \in \mathcal{D}} \frac{|g(x) - g(y)|}{\|x - y\|}.$$

Let P_n be the empirical distribution associated with $x_1, \ldots, x_n$, i.e. $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. We begin by providing a deterministic bound involving the distance between the measures P_n and P. As these are measures over different spaces, it is not obvious how to relate them. We will use transportation distance to do so. Recall that a mapping $T : \mathcal{D} \to \mathcal{D}$ is a *transportation map* between P and P_n if for all measurable $A \subseteq \mathcal{D}$,

$$P_n(A) = P(T^{-1}(A))$$

Lemma 2. *For any $f \in C^{0,1}_{\mathcal{D}}$, the following bound holds on the Sobolev discrete norm:*

$$\|B\beta\|_2^2 \leq n^2 r^2 (r + \|\mathrm{Id} - T\|_{L^\infty(P)})^d$$

Proof. We write

$$\begin{aligned}
\frac{1}{n^2}\|B^{(s)}\beta\|_2^2 &= \frac{1}{n^2}\beta^T L \beta \\
&= \frac{1}{2n^2}\sum_{i,j=1}^{n}(\beta_i - \beta_j)^2 A_{ij} \\
&= \int_{\mathcal{D}}\int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x,y)\, dP_n(y)\, dP_n(x)
\end{aligned}$$

We examine the inner integral. By the Holder property of f, and a change of variables, we obtain

$$\begin{aligned}
\int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x,y)\, dP_n(y) &\leq \int_{\mathcal{D}} \|x - y\|^2 \eta_r(x,y)\, dP_n(y) \\
&= \int_{\mathcal{D}} \|x - T(y)\|^2 \eta_r(x, T(y))\, dP(y) \\
&\leq r^2 \int_{\mathcal{D}} \eta_r(x, T(y))\, dP(y) \\
&\leq r^2 \big(r + \|\mathrm{Id} - T\|_{L^\infty(P)}\big)^d
\end{aligned}$$

and the desired result is shown. □

2.2 $s = 1$, $f \in \mathcal{W}^{1,2}(\mathcal{D}; 1)$.

Let P be an absolutely continuous probability measure over $\mathcal{D}$ with density function p bounded above and below by constants, i.e

$$0 < p_{\min} < p(x) < p_{\max} < \infty, \quad \text{for all } x \in \mathcal{D}$$

Lemma 3. *For any $f \in W^{1,2}(\mathcal{D}; L)$, and any $b \geq 1$, we have that there exists a constant $c_2 > 0$ which depends only on d and $p_{\max}$ such that*

$$\|B\beta\|_2^2 \leq L^2 b^2 c_2 n^2 r^{d+2} \tag{1}$$

with probability at least $1 - \frac{1}{b^2}$.

Proof. Observe that

$$\begin{aligned}
\frac{1}{n^2}\mathbb{E}(\beta^T L \beta) = \mathbb{E}\left(\frac{1}{n^2}\sum_{i,j=1}^{n}(\beta_i - \beta_j)^2 A_{ij}\right) &= \frac{(n-1)}{n}\int_{\mathcal{D}}\int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x,y)\, dP(y)\, dP(x) \\
&\leq p_{\max}^2 \frac{(n-1)}{n}\int_{\mathcal{D}}\int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x,y)\, dy\, dx.
\end{aligned}$$

We will show that for any $f \in \mathcal{W}^{1,2}(\mathcal{D}; L)$, there exists a constant c_4 which depends only on dimension d such that

$$\int_{\mathcal{D}} \int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x, y) \, dy \, dx \leq c_4 L^2 r^{d+2} \tag{2}$$

whence the desired result of (1) follows by Markov's inequality.

We begin by dealing with complications due to the boundary of $\mathcal{D}$. Let V be any bounded open set such that $\mathcal{D} \subset\subset V$. Note that as $\partial\mathcal{D}$ is C^1, by Theorem 1 there exists $g \in W^{1,2}(\mathbb{R}^d)$ such that

1. $g = f$, P-almost-everywhere in $\mathcal{D}$
2. g has support within V, and
3. $\|g\|_{W^{1,2}(\mathbb{R}^d)} \leq C \|f\|_{W^{1,2}(\mathcal{D})}$ for a constant C which depends only on $\mathcal{D}$.

As a result of the first point,

$$\int_{\mathcal{D}} \int_{\mathcal{D}} (f(x) - f(y))^2 \eta_r(x, y) \, dy \, dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (g(x) - g(y))^2 \eta_r(x, y) \, dy \, dx. \tag{3}$$

Next, we smooth g, so that we may work with ordinary partial derivatives. We let $\kappa \in C^\infty(\mathbb{R}^d)$ be given by

$$\kappa(x) := \begin{cases} C \exp\left\{\frac{1}{\|x\|^2 - 1}\right\} & \text{if } \|x\|_2 \leq 1 \\ 0 & \text{if } \|x\|_2 \geq 1 \end{cases}$$

where the normalizing constant $C > 0$ is chosen so that $\int_{\mathbb{R}^d} \eta dx = 1$. Let $\kappa_r(x) := (1/r^d)\kappa(x/r)$. Then, the mollification of g by κ_r is given by

$$\begin{aligned} g^r &:= g * \eta_r \\ &= \int_{\mathbb{R}^d} \eta_r(x - y) g(y) dy \end{aligned}$$

(Refer to [Evans, 2010], Appendix C, Theorem 7 for a proof that $g^r \in C^\infty(\mathbb{R}^d)$.) Adding and subtracting within (3), we have

$$\begin{aligned} &\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (g(x) - g(y))^2 \eta_r(x, y) \, dy \, dx & (4) \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big((g(x) - g^r(x))^2 + (g^r(x) - g^r(y))^2 + (g^r(y) - g(y))^2\big) \eta_r(x, y) \, dy \, dx \\ &= 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big((g(y) - g^r(y))^2 \eta_r(x, y) \, dy \, dx + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big((g^r(x) - g^r(y))^2 \eta_r(x, y) \, dy \, dx & (5) \end{aligned}$$

We deal with each summand individually, beginning with the first one. We have

$$\begin{aligned} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big((g(y) - g^r(y))^2 \eta_r(x, y) \, dy \, dx &= \int_{\mathbb{R}^d} \int_{B(x,r)} \big((g(y) - g^r(y))^2 \, dy \, dx \\ &= \int_{\mathbb{R}^d} \int_{B(y,r)} \big((g(y) - g^r(y))^2 \, dx \, dy & \text{(Tonelli's Theorem)} \\ &= r^d \int_{\mathbb{R}^d} \big((g(y) - g^r(y))^2 \, dy & (6) \\ &\leq r^{d+2} \int_{\mathbb{R}^d} \|\nabla g(y)\|^2 \, dy & (7) \end{aligned}$$

where the last line follows from Lemma 4.

We now turn our attention to the second summand. Note that as $g^r \in C^\infty(\mathbb{R}^d)$, we may apply Theorem 3 and obtain

$$\begin{aligned}(g^r(x) - g^r(y))^2 &= \left(\int_0^1 \nabla g^r(x + t(y-x)) \cdot (y-x)\, dt\right)^2 \\ &\le \int_0^1 \left(\nabla g^r(x+t(y-x)) \cdot (y-x)\right)^2 dt && \text{(Jensen's inequality)} \\ &\le \|y-x\|^2 \int_0^1 \|\nabla g^r(x+t(y-x))\|^2\, dt. && \text{(Cauchy-Schwarz inequality)}\end{aligned}$$

As a result, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left((g^r(x)-g^r(y))^2 \eta_r(x,y)\, dy\right) dx &\le \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \|y-x\|^2 \int_0^1 \|\nabla g^r(x+t(y-x))\|^2 \eta_r(x,y)\, dt\, dy\, dx \\ &\le r^2 \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_0^1 \|\nabla g^r(x+t(y-x))\|^2 \eta_r(x,y)\, dt\, dy\, dx \\ &= r^2 \int_{\mathbb{R}^d}\int_0^1\int_{\mathbb{R}^d} \|\nabla g^r(x+t(y-x))\|^2 \eta_r(x,y)\, dx\, dt\, dy \\ &= r^2 \int_{\mathbb{R}^d}\int_0^1\int_{\mathbb{R}^d} \|\nabla g^r(x+tz)\|^2 \eta_r(z)\, dx\, dt\, dz && (z = y - x)\end{aligned}$$

where we write $\eta_r(z) = \mathbf{1}(\|z\| \le r)$ in an abuse of notation. Next, we note that

$$\int_{\mathbb{R}^d} \|\nabla g^r(x+tz)\|^2\, dx = \int_{\mathbb{R}^d} \|\nabla g^r(x)\|^2\, dx \le \int_{\mathbb{R}^d} \|\nabla g(x)\|^2\, dx$$

with the inequality following from Lemma 5. Therefore,

$$\begin{aligned}r^2 \int_{\mathbb{R}^d}\int_0^1\int_{\mathbb{R}^d} \|\nabla g^r(x+tz)\|^2 \eta_r(z)\, dx\, dt\, dz &\le r^2 \int_{\mathbb{R}^d} \eta_r(z) \int_0^1 \int_{\mathbb{R}^d} \|\nabla g(x)\|^2\, dx\, dt\, dz \\ &= r^{2+d} \int_{\mathbb{R}^d} \|\nabla g(x)\|^2\, dx. && (8)\end{aligned}$$

By (5), (7) and (8), we have that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (g(x)-g(y))^2 \eta_r(x,y)\, dx\, dy \le 3 \int_{\mathbb{R}^d} \|\nabla g(x)\|_2^2\, dx$$

Then by (3), $\int_{\mathbb{R}^d} \|\nabla g(x)\|_2^2\, dx \le C \int_{\mathcal{D}} \|\nabla f(x)\|_2^2\, dx \le C$ where C is a constant depending only on $\mathcal{D}$. So the desired result of (3) follows. □

3 Supporting Results.

Theorem 1 ([Evans, 2010] Chapter 5.4, Theorem 1)**.** *Assume U is bounded and ∂U is C^1. Select a bounded open set V such that $U \subset\subset V$ (U is compactly contained in V). Then there exists a bounded linear operator $E : W^{1,2}(U) \to W^{1,2}(\mathbb{R}^d)$ such that for each $u \in W^{1,2}(U)$:*

1. *$Eu = u$ a.e. in U,*
2. *Eu has support within V, and*

3.

$$\|Eu\|_{W^{1,2}(\mathbb{R}^d)} \le C\|u\|_{W^{1,2}(\mathbb{R}^d)}$$

the constant C depending only on U and V.

For $u \in W^{1,2}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, write $\nabla u(x) = (D^{e_1}(x), \dots, D^{e_d}(x)$, for the gradient of u.

Theorem 2 ([Evans, 2010] Chapter 5.8.1, Theorem 2)**.** *There exists a constant C, depending only on d, such that*

$$\|u - u^r\|_{L^2(B(x,r))} \le Cr\|\nabla u\|_{L^2(B(x,r))}$$

Theorem 3 (Taylor expansion.)**.** *For any function $u \in C^1$, and any $x, y \in \mathbb{R}^d$,*

$$u(y) - u(x) = \int_0^1 \nabla(u(x + t(y-x))) \cdot (y-x)\, dt$$

Lemma 4. *For any function $g \in W^{1,2}(\mathbb{R}^d)$ compactly supported in a bounded open set $V \subset \mathbb{R}^d$, we have*

$$\int_{\mathbb{R}^d} (g(x) - g^r(x))^2\, dx \le r^2 \int_{\mathbb{R}^d} \|\nabla g(x)\|^2\, dx \tag{9}$$

Proof. This Lemma is essentially a reproduction of part of the proof of the Rellich-Kondrachov Compactness Theorem from [Evans, 2010]. Note that it is sufficient to prove in the case when g is smooth. To see this, for the moment assume (9) holds for all $u \in C^\infty(V)$, and let $g \in W^{1,2}(V)$. By Theorem 4, we may take a sequence $g_m \in C^\infty(V)$ such that

$$\|g - g_m\|_{L^2(V)} \to 0, \quad \text{and} \quad \|\nabla g - \nabla g_m\|_{L^2(V)} \to 0,$$

Then we have

$$\begin{aligned}\int_{\mathbb{R}^d} (g(x) - g^r(x))^2\, dx &\le \int_{\mathbb{R}^d} (g(x) - g_m(x))^2\, dx + \int_{\mathbb{R}^d} (g_m(x) - g_m^r(x))^2\, dx + \int_{\mathbb{R}^d} (g_m^r(x) - g^r(x))^2\, dx \\ &\le \int_{\mathbb{R}^d} (g(x) - g_m(x))^2\, dx + r^2 \int_{\mathbb{R}^d} \|\nabla g_m(x)\|^2\, dx + \int_{\mathbb{R}^d} (g_m^r(x) - g^r(x))^2\, dx\end{aligned}$$

and taking the limit as m goes to infinity, the right hand side converges to $\int_{\mathbb{R}^d} \|\nabla g(x)\|^2\, dx$.

It remains to show (9) in the case where g is smooth. In this case, we have

$$\begin{aligned}g^r(x) - g(x) &= \frac{1}{r^d} \int_{B(x,r)} \kappa\left(\frac{x-z}{r}\right) \big(g(z) - g(x)\big)\, dz \\ &= \int_{B(0,1)} \kappa(y) \big(g(x - ry) - g(x)\big)\, dy \\ &= \int_{B(0,1)} \kappa(y) \int_0^1 \frac{d}{dt}\big(g(x - try)\big)\, dt\, dy \\ &= -r \int_{B(0,1)} \kappa(y) \int_0^1 \big(\nabla g(x - try)\big) \cdot y\, dt\, dy.\end{aligned}$$

Therefore, by Jensen's and Cauchy-Schwarz inequalities, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} (g(x) - g^r(x))^2 \, dx &\leq r^2 \int_{\mathbb{R}^d} \int_{B(0,1)} \kappa(y) \int_0^1 \|\nabla g(x - try)\|^2 \|y\|^2 \, dt \, dy \, dx \\
&\leq r^2 \int_{B(0,1)} \kappa(y) \int_0^1 \int_{\mathbb{R}^d} \|\nabla g(x - try)\|^2 \, dx \, dt \, dy \\
&= r^2 \int_{\mathbb{R}^d} \|\nabla g(z)\|^2 \, dz
\end{aligned}$$

□

The following theorem is Theorem 1 in Section 5.3 of [Evans, 2010].

Theorem 4 (Local approximation by smooth functions.)**.** *Assume U is bounded, and $u \in W^{k,p}(U)$ for some $1 \leq p < \infty$. Then there exists functions $u_m \in C^\infty(U) \cap W^{k,p}(U)$ such that*

$$\|u_m - u\|_{W^{k,p}(U)} \overset{m}{\to} 0.$$

Lemma 5. *For any $u \in W^{1,2}(\mathbb{R}^d)$, we have*

$$\int_{\mathbb{R}^d} \|\nabla u^r(x)\|_2^2 \, dx \leq \int_{\mathbb{R}^d} \|\nabla u(x)\|_2^2 \, dx$$

Proof. Observe that $\|\nabla u(x)\|_2^2 = \sum_{j=1}^d (D^{e_j} u(x))^2$. Therefore,

$$\begin{aligned}
\int_{\mathbb{R}^d} \|\nabla u^r(x)\|_2^2 \, dP(x) &= \sum_{j=1}^d \int_{\mathbb{R}^d} (D^{e_j} u^r(x))^2 \, dx \\
&= \sum_{j=1}^d \int_{\mathbb{R}^d} ((D^{e_j} u)^r(x))^2 \, dx
\end{aligned}$$

where the second equality follows from equation (1) in Section 5.3 of [Evans, 2010]. Then, for any $v \in L^2(\mathbb{R}^d)$, we have that

$$\begin{aligned}
|v^r(x)| &= \int_{\mathbb{R}^d} \kappa_r^{1/2}(x-y) \kappa_r^{1/2}(x-y) v(y) \, dy \\
&\leq \left(\int_{\mathbb{R}^d} \kappa_r^(x - y) \, dy \right)^{1/2} \left(\int_{\mathbb{R}^d} \kappa_r(x-y) v^2(y) \right)^{1/2} \\
&= \left(\int_{\mathbb{R}^d} \kappa_r(x-y) v^2(y) \right)^{1/2}
\end{aligned}$$

and therefore

$$\begin{aligned}
\int_{\mathbb{R}^d} (v^r(x))^2 \, dx &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_r(x-y) v^2(y) \, dy \, dx \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_r(x-y) v^2(y) \, dy \, dx \\
&= \int_{\mathbb{R}^d} v^2(y) \, dy
\end{aligned}$$

Applying this to $D^{e_j}u \in L^2(\mathbb{R}^d)$, we have that

$$\sum_{j=1}^{d} \int_{\mathbb{R}^d} ((D^{e_j}u)^r(x))^2 \, dx \leq \sum_{j=1}^{d} \int_{\mathbb{R}^d} ((D^{e_j}u(x))^2 \, dx = \int_{\mathbb{R}^d} \|\nabla u(x)\|^2 \, dx.$$

$\square$

REFERENCES

Lawrence C. Evans. *Partial differential equations.* American Mathematical Society, 2010.