Supplementary material for the paper: Robust Nonparametric Regression with Metric-Space valued Output

Matthias Hein
Department of Computer Science, Saarland University
Campus E1 1, 66123 Saarbrücken, Germany
hein@cs.uni-sb.de

1 Missing Proof from Section 2

Lemma 1 Let \( N \) be a complete metric space such that \( d(x,y) < \infty \) for all \( x, y \in N \) and every closed and bounded set is compact. If \( \Gamma \) is \((\alpha, s)\)-bounded and \( \Gamma'_1(x, q) < \infty \) for some \( q \in N \), then

- \( R'_1(x, p) \) is continuous for all \( p \in N \),
- \( R'_1(x, \cdot) \) is continuous on \( N \),
- The set of minimizers \( Q^* = \arg\min_{q\in N} R'_1(x, q) \) exists and is compact.

Proof: As \( \Gamma \) is monotonically increasing and convex, we have for any \( p, y \in N \),

\[
\Gamma(d_N(p,y)) \leq \Gamma(d_N(p,q) + d_N(q,y)) \leq \frac{1}{2}[\Gamma(2d_N(p,q)) + \Gamma(2d_N(q,y))],
\]

Moreover, since \( \Gamma \) is \((\alpha, s)\)-bounded we have,

\[
\Gamma(2x) \leq a\Gamma(x)\mathbb{1}_{x \geq s} + \Gamma(2s)\mathbb{1}_{x < s}.
\]

Taking expectations with respect to \( Y|X = x \) we get,

\[
R'_1(x, p) \leq \Gamma(2s) + \frac{a}{2}\Gamma(d_N(p,q)) + \frac{a}{2}R'_1(x, q).
\]

Next, we show continuity of \( R'_1(x, \cdot) \). Using Lemma 2 we get,

\[
|R'_1(x, p) - R'_1(x, q)| = |E[\Gamma(d_N(p, Y)) - \Gamma(d_N(q, Y))]| \leq d(p,q)E[\max\{\Gamma'(d_N(p, Y)), \Gamma'(d_N(q, Y))\}]|.
\]

Now, for \( x \geq s \) we have \( \Gamma'(x) \leq \frac{\Gamma(2x) - \Gamma(x)}{x} \leq (a - 1)\frac{\Gamma(x)}{x} \) and for \( x < s \), \( \Gamma'(x) \leq \Gamma'(s) \). Thus

\[
E[\Gamma'(d_N(p, Y))] \leq \frac{(a-1)}{s}E[\Gamma(d_N(p, Y))] + \Gamma'(s),
\]

which shows using \( \max\{a, b\} \leq a + b \) the continuity of \( R'_1(x, \cdot) \).

Finally, we consider the set \( S_\varepsilon = \{ q \in N \mid R'_1(x, q) \leq \inf_{p \in N} R'_1(x, p) + \varepsilon \} \) which is closed since \( R'_1(x, \cdot) \) is continuous. Moreover, let \( q_1, q_2 \in S_\varepsilon \), then

\[
\Gamma(d_N(q_1, q_2)) \leq \Gamma(2s) + \frac{a}{2}\Gamma(d_N(q_1, y)) + \frac{a}{2}\Gamma(d_N(q_2, y)) \leq \Gamma(2s) + \frac{a}{2}R'_1(x, q_1) + \frac{a}{2}R'_1(x, q_2).
\]

For \( x \geq s \) we have shown above \( x \leq (a - 1)\frac{\Gamma(x)}{\Gamma'(x)} \leq \frac{\Gamma(x)}{\Gamma'(s)} \) and thus either \( d_N(q_1, q_2) \leq s \) or

\[
d_N(q_1, q_2) \leq (a - 1)\frac{\Gamma(2s) + \frac{a}{2}R'_1(x, q_1) + \frac{a}{2}R'_1(x, q_2)}{\Gamma'(s)}.
\]
which shows that the set \( S_\varepsilon \) is bounded and thus compact. It is non-empty since \( R_t^r(x, \cdot) \) is continuous. The set of minimizers \( Q^* = \cap_{\varepsilon > 0} S_\varepsilon \) is compact and non-empty as it is the intersection of a nested sequence of non-empty, compact sets.

\[ \square \]

## 2 Missing Proofs from Section 5 and 7

The supplementary material contains the proofs which due to space constraints could not be included into the paper. For convenience we restate here Assumptions (A1) from the paper.

**Assumptions (A1):**

- \((X_t, Y_t)_{t=1}^T\) is an i.i.d. sample of \( P \) on \( M \times N \),
- \( M \) and \( N \) are compact manifolds,
- The data-generating measure \( P \) on \( M \times N \) is absolutely continuous with respect to the natural volume element,
- The marginal density on \( M \) fulfills: \( p(x) \geq p_{\min}, \forall x \in M \),
- The density \( p(y, \cdot) \) is continuous on \( M \) for all \( y \in N \),
- The kernel fulfills: \( a s \leq k(s) \leq b e^{-\gamma s^2} \) and \( \int_{\mathbb{R}^m} ||x|| k(||x||) \, dx < \infty \),
- The loss \( \Gamma : \mathbb{R}_+ \to \mathbb{R}_+ \) is \((\alpha, s)\)-bounded.

This proposition collects results from [1].

**Proposition 1** Let \( M \) be a compact \( m \)-dimensional Riemannian manifold. Then, there exists \( r_0 > 0 \) and \( S_1, S_2 > 0 \) such that for all \( x \in M \) all balls \( B(x, r) \) with radius \( r \leq r_0 \) it holds,

\[
S_1 \, r^m \leq \text{vol} \left( B(x, r) \right) \leq S_2 \, r^m.
\]

Moreover, the cardinality \( K \) of a \( \delta \)-covering of \( M \) is upper bounded as, \( K \leq \frac{\text{vol}(N)}{S_1} \left( \frac{2}{\delta} \right)^m \).

**Proposition 2** Let the assumptions A1 hold, then if \( f \) is continuous we get for any \( x \in M \setminus \partial M \),

\[
\lim_{h \to 0} \int_M k_h(d_M(x, z)) f(z) \, dV(z) = C_x f(x),
\]

where \( C_x = \lim_{h \to 0} \int_M k_h(d_M(x, z)) \, dV(z) > 0 \). If moreover \( f \) is Lipschitz continuous with Lipschitz constant \( L \), then there exists a \( h_0 > 0 \) such that for all \( h < h_0(x) \),

\[
\int_M k_h(d_M(x, z)) f(z) \, dV(z) = C_x f(x) + O(h).
\]

**Proof:** We denote by \( \text{inj}(M) \) the injectivity radius of \( M \). As \( f \) is continuous for any \( \varepsilon > 0 \), \( \exists \delta \) such that \( d(x, z) < \delta \) implies \( |f(x) - f(z)| < \varepsilon \). Suppose that \( \varepsilon \) is chosen small enough, so that \( \delta < \text{inj}(M) \),

\[
\int_M k_h(d_M(x, z)) (f(z) - f(x)) \, dV(z) \\
\leq \varepsilon \int_{B(x, \delta)} k_h(d_M(x, z)) \, dV(z) + 2 \|f\|_\infty \int_{M \setminus B(x, \delta)} k_h(d_M(x, z)) \, dV(z) \\
\leq \varepsilon \int_{B(x, \delta)} k_h(||y||) \, dy + \|f\|_\infty \frac{\text{vol}(M)}{h^m} b e^{-\gamma \frac{s^2}{2}},
\]

where we have introduced in the last step normal coordinates centered at \( x \) on \( B(x, \delta) \) so that \( d_M(x, z) = ||y|| \). Note, that the second term is independent of \( \varepsilon \) and for each \( \delta > 0 \) converges
Lemma 2

Let \( f \) be convex, differentiable and monotonically increasing. Then
\[
\min \{ \phi'(x), \phi'(y) \} |y - x| \leq |\phi(y) - \phi(x)| \leq \max \{ \phi'(x), \phi'(y) \} |y - x|.
\]

Proof: Using the first order condition of a convex function and \( \phi(x) \leq \phi(y) \) for \( x \leq y \),
\[
\phi(y) - \phi(x) \geq \phi'(x)(y - x) \quad \Rightarrow \phi(x) - \phi(y) \leq \phi'(x)(x - y),
\]
\[
\phi(x) - \phi(y) \geq \phi'(y)(y - x) \quad \Rightarrow \phi(y) - \phi(x) \leq \phi'(y)(y - x).
\]
The left part yields the lower bound and the right part the upper bound.

References