A Appendix

A.1 Extending the pointwise normal reach to the non-manifold setting

Federer (1959) introduces reach for arbitrary subsets of Euclidean space. In this situation $T_{\mathbf{x}}\mathcal{M}$ and $N_{\mathbf{x}}\mathcal{M}$ denote the tangent- and normal cone.

Definition A.1. Let $\mathcal{M} \subset \mathbb{R}^D$ denote an arbitrary subset and let $x \in \mathcal{M}$. Then $v \in \mathbb{R}^D$ is a tangent vector for $\mathcal{M}$ at $\mathbf{x}$ if either $v = 0$ or if for every $\varepsilon > 0$ exists $\mathbf{y} \in \mathcal{M}$ with

$$0 < \|\mathbf{y} - \mathbf{x}\| < \varepsilon \quad \text{and} \quad \left\| \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|} - \frac{v}{\|v\|} \right\| < \varepsilon. \tag{23}$$

Let $T_{\mathbf{x}}\mathcal{M}$ denote the set of tangent vectors for $\mathcal{M}$ at $\mathbf{x}$. A vector $w \in \mathbb{R}^D$ is a normal vector for $\mathcal{M}$ at $\mathbf{x}$ if

$$\langle w, v \rangle \leq 0 \text{ for all } v \in T_{\mathbf{x}}\mathcal{M}. \tag{24}$$

Let $N_{\mathbf{x}}\mathcal{M}$ denote the set of all normal vectors for $\mathcal{M}$ at $\mathbf{x}$.

We can extend theorem 2.4 and Lemma 2.6 to the general situation as defined by Federer. To extend Theorem 2.4 it is sufficient to prove that for any $v \in N_{\mathbf{x}}\mathcal{M}$ and $u \in \mathbb{R}^D$, $\|P_v(u)\| \leq d(u, T_{\mathbf{x}}\mathcal{M})$.

Lemma A.2. *For any $v \in N_{\mathbf{x}}\mathcal{M}$ and $u \in \mathbb{R}^D$ with $\langle v, u \rangle \geq 0$, $\|P_v(u)\| \leq d(u, T_{\mathbf{x}}\mathcal{M})$.*

Proof. For a subset $A \subset \mathbb{R}^D$, $\mathrm{dual}(A) = \{v \in \mathbb{R}^D : \langle a, v \rangle \leq 0 \text{ for all } a \in A\}$. First we prove that $d(u, \mathrm{dual}(v)) = \|P_v(u)\|$. Note that we can write $u = u_v + u_{v^\perp}$, where $u_v = P_v(u)$ and $u_{v^\perp} \in v^\perp$. Then

$$\begin{aligned} d(u, \mathrm{dual}(v)) &= \inf_{w \in \mathrm{dual}(v)} \|u - w\| = \inf_{w \in \mathrm{dual}(v)} \|u_v + u_{v^\perp} - w\| \\ &= \inf_{w \in \mathrm{dual}(v)} \|u_v\| + \|u_{v^\perp} - w\| - 2 \langle u_v, w \rangle . \end{aligned} \tag{25}$$

As $\langle u_v, w \rangle \leq 0$, it follows that the infimum is achieved when $w = u_{v^\perp}$. By the definition of the dual it follows that $\mathrm{dual}(v) \supset \mathrm{dual}(N_{\mathbf{x}}\mathcal{M}) \supset T_{\mathbf{x}}\mathcal{M}$. Hence $d(u, T_{\mathbf{x}}\mathcal{M}) = \inf_{w \in T_{\mathbf{x}}\mathcal{M}} \|u - w\| \geq \inf_{w \in \mathrm{dual}(v)} \|u - w\| = \|P_u(v)\|$. ∎

To extend lemma 2.6 note that if $r_{max}(x) > 0$, then $T_{\mathbf{x}}\mathcal{M}$ is convex (Federer, 1959, Thm 4.8 (12)). Let $\mathbf{y} \in \mathcal{M}$. If $\mathbf{y} - \mathbf{x} \in T_{\mathbf{x}}\mathcal{M}$ then $R(\mathbf{x}, \mathbf{y}) = \infty$. Otherwise, as $T_{\mathbf{x}}\mathcal{M}$ is a convex cone, there exists $n \in N_{\mathbf{x}}\mathcal{M}$ such that $\langle n, \mathbf{y} - \mathbf{x} \rangle \geq 0$. In that case $\langle n, \mathbf{x} - \mathbf{y} \rangle = \|P_n(\mathbf{y} - \mathbf{x})\|$, so applying Lemma 2.5 gives the result.

Though the theory can be extended to general subspaces, the manifold assumption is important for the experimental setup. An important assumption for the estimator (16) is that the Jacobian spans the entire tangent space. If this is not the case, this estimator does not estimate the pointwise normal reach. The reason being, the length of the projection onto the orthogonal complement of the Jacobian is not necessarily the distance to the tangent space. It is clear that when we want to study the uniqueness of latent representations, if the decoder is not injective, it automatically has areas without unique representations. So if the decoder is not injective, we should already be wary about trusting the latent representations.