1 Continuing from Lecture 23

Here we resume the proof of Lemma 3 from Lecture 23. Recall that we wanted to establish a bound
\[ \text{Vol}(B_r(p) \cap M) \geq f(r, \tau) \] for “most” \( p \in M \)
for some function \( f \) of \( r \) and \( \tau := \text{reach}(M) \).

Proof. We saw that if
\[ B_{r'}(p) \cap T_p \subseteq \Pi_{T_p}(B_r(p) \cap M) \] (1)
for some
\[ r' \geq \sqrt{1 - \frac{r^2}{4\tau^2}} \cdot r, \]
then we would obtain the desired result.

Now, from Lemma 1 of Lecture 23, we know that \( \Pi_{T_p} \) is invertible on \( B_r(p) \cap M \) for all \( r \in [0, \frac{\tau}{\sqrt{2}}) \). This fact implies that \( \Pi_{T_p}(B_r(p) \cap M) \) is open in \( T_p \). Thus, there exists \( s \in \mathbb{R}^+ \) such that
\[ B_s(p) \cap T_p \subseteq \Pi_{T_p}(B_r(p) \cap M). \] (2)

Let \( s^* \) be the supremum of all \( s \in \mathbb{R}^+ \) satisfying (2).

There is \( y \in \partial(B_{s^*}(p) \cap T_p) \cap \partial(\Pi_{T_p}(B_r(p) \cap M)) \). Set
\[ x := \Pi_{T_p}^{-1}(y). \]

One can see that \( x \in \partial(B_r(p) \cap M) \). Hence
\[ \|x - p\|_2 = r \]
as long as, e.g., \( B_r(p) \cap \partial M = \emptyset \). Finally, set
\[ t := \|x - y\|_2. \]
Lemma 2 of Lecture 23 tells us

\[ \angle yp x \leq \arcsin \left( \frac{r}{2\tau} \right), \]

implying

\[ \frac{t}{r} \leq \frac{r}{2\tau}. \]

And so

\[ s^* \geq \sqrt{1 - \frac{r^2}{4\tau^2}} \cdot r, \]

meaning (1) follows by setting \( r' := s^* \).

The discussion at the end of Lecture 22 now gives us a covering number bound for at least the interior of \( \mathcal{M} \).

**Definition 1.** For a \( d \)-dimensional manifold \( \mathcal{M} \subseteq \mathbb{R}^D \), the \textbf{r-interior} of \( \mathcal{M} \) is

\[ \text{int}_r(\mathcal{M}) \overset{\text{def}}{=} \{ p \in \mathcal{M} : B_r(p) \cap \partial \mathcal{M} = \emptyset \} \]

We have proven the following result, which will help us prove Theorem 2, the desired manifold embedding result.

**Theorem 1.** Let \( \mathcal{M} \subseteq \mathbb{R}^D \) be a \( d \)-dimensional manifold with \( \tau := \text{reach}(\mathcal{M}) > 0 \). Let \( r \in [0, \frac{\tau}{4}] \). Then the covering number will obey

\[ C_r(\text{int}_r(\mathcal{M})) \leq \frac{\text{Vol}_d(\mathcal{M}) \left( 1 - \frac{r^2}{4\tau^2} \right)^{-\frac{d}{2}}} {\text{Vol}(\text{unit ball in } \mathbb{R}^d)}. \]
2 The Johnson-Lindenstrauss Lemma and Manifold Embeddings

We wanted to show that a random matrix (in our case, one with subgaussian entries) will nearly isometrically embed any compact, $d$-dimensional manifold $M \subseteq \mathbb{R}^D$ with positive reach, into $\mathbb{R}^m$ such that $m \ll D$. The following theorem tells us precisely what this means.

**Theorem 2.** Let $M \subseteq \mathbb{R}^D$ be a $d$-dimensional, $C^2$-manifold with $\text{Vol}_d(M) < \infty$, $\tau := \text{reach}(M) > 0$, and

$$d(p, \text{int}_r(M)) \leq r \text{ for all } r \in \left[0, \frac{\tau}{4}\right], \text{ for all } p \in M.$$  

Let $\epsilon, \delta \in (0, 1)$. Finally, let $A \in \mathbb{R}^{m \times D}$ with i.i.d. subgaussian entries (with parameter $c$). Then

$$-\delta + (1 - \epsilon)\|x - y\|_2 \leq \left\| \frac{1}{\sqrt{m}}A(x - y) \right\|_2 \leq (1 + \epsilon)\|x - y\|_2 + \delta \text{ for all } x, y \in M$$

with probability at least $p \in (0, 1)$, provided

$$m \geq \frac{(64c)(16c + 1)}{\epsilon^2} \ln \left( \frac{8}{1 - p} C_{\tilde{r}}^2(\text{int}_{\tilde{r}}(M)) \right),$$

for

$$\tilde{r} := \min \left\{ \sqrt{\frac{d}{D}} \frac{\delta}{18 \sqrt{\epsilon}}, \frac{\tau}{4} \right\}.$$

• Theorem 1 tells us

$$C_{\tilde{r}}(\text{int}_{\tilde{r}}(M)) \leq \frac{\text{Vol}_d(M)}{\text{Vol}(\text{unit ball in } \mathbb{R}^d)} \left( \frac{16}{15} \right)^{\frac{d}{2}} \max \left\{ \sqrt{\frac{D}{d}} \cdot \frac{18 \sqrt{\epsilon}}{\delta \tau}, \frac{4}{\tau} \right\}^d.$$

Thus,

$$m \sim C' \frac{d}{\epsilon^2} \ln \left( \frac{\tilde{C}}{\min\{\tau, 1\}(1 - p)\delta} \cdot \frac{D}{d} \right)$$

for $D > 2d$ and constants $C'$ and $\tilde{C}$ depending on $c$, and log $(\text{Vol}_d(M))$, and assuming $d \ll D$.

With a bit more work, one can prove variants of Theorem 1 that make $m$ independent of $D$, specifically

$$m \sim C' \frac{d}{\epsilon^2} \ln \left( \frac{\tilde{C}d}{\min\{\tau, 1\}(1 - p)} \right).$$

With a substantial amount of work, one can prove

$$m \sim C' \frac{d}{\epsilon^2} \ln \left( \frac{\tilde{C}d}{\min\{\tau, 1\}(1 - p)} \right).$$
For these results, see [1, 2], respectively.

Let’s now prove Theorem 2.

Proof. Let $C \subseteq \mathcal{M}$ be a minimal $\tilde{r}$-cover of $\text{int}_{\tilde{r}}(\mathcal{M})$. Note that $C$ is also a $2\tilde{r}$-cover of $\mathcal{M} \subseteq \mathbb{R}^D$.

Theorem 1 of Lecture 14 guarantees that $\tilde{A} := \frac{1}{\sqrt{m}}A$, with $m$ as above, will satisfy

$$(1 - \epsilon) \leq \sqrt{1 - \epsilon} \leq \frac{\|A(p - q)\|_2}{\|p - q\|_2} \leq \sqrt{1 + \epsilon} \leq 1 + \epsilon \quad \text{for all } p, q \in C$$

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

Now, Theorem 1 of Lecture 15 guarantees that $\tilde{A}$ also has the RIP of order $d$ for $\epsilon < 1$, with probability at least $1 - \frac{1 - p^2}{2}$. That is, $\epsilon_d(A) \in (0, 1)$, implying

$$\sigma_1(\tilde{A}) \leq 2\sqrt{2} \sqrt{\frac{D}{d}}$$

by Lemma 2 of Lecture 16. The union bound implies that (3) and (4) hold simultaneously with probability at least $p$.

Thus,

$$\|\tilde{A}(x - y)\|_2 \leq \|\tilde{A}(x - p_x)\|_2 + \|\tilde{A}(p_x - p_y)\|_2 + \|\tilde{A}(p_y - y)\|_2$$

$$\leq 2\sqrt{2} \sqrt{\frac{D}{d}} (\|x - p_x\|_2 + \|p_y - y\|_2) + (1 + \epsilon)\|p_x - p_y\|_2,$$

where $p_x$ and $p_y$ are the closest points in $C$ to $x$ and $y$, respectively. That is,

$$p_x = \arg \min_{p \in C} \|p - x\|_2 \quad \text{and} \quad p_y = \arg \min_{p \in C} \|p - y\|_2.$$

As $C$ is a $2\tilde{r}$-cover, $\|x - p_x\|_2$ and $\|p_y - y\|_2$ are bounded from above by $2\tilde{r}$, while an additional application of the triangle inequality gives $\|p_x - p_y\|_2 \leq \|x - y\|_2 + 4\tilde{r}$. When used above, these estimates yield

$$\|\tilde{A}(x - y)\|_2 \leq \frac{6\delta}{9} + (1 + \epsilon)\|x - y\|_2,$$

giving the desired upper bound. An analogous argument gives the desired lower bound. \qed

References
