Supplementary material - Bilinear Random Projections for Locality-Sensitive Binary Codes

April 16, 2015

1 Outline

In this supplementary material, we provide the proof of Lemma 3 and Theorem 3 described in the paper. Note that Lemma 3 and Theorem 3 in the paper are renamed as Lemma 2 and Theorem 1 in this supplementary material, respectively.

2 Bounds on the Expected Hamming Distance

Lemma 1 *[1] For any* $u, v \in [-1, 1]$, $P_t\{sgn(u+t) \neq sgn(v+t)\} = |u-v|/2$.

Lemma 2 *(Lemma 2 in the paper)*

$$\mathbb{E}_{\boldsymbol{w},\boldsymbol{v},b,t}\left[\mathbb{I}_{h(\boldsymbol{X})\neq h(\boldsymbol{Y})}\right] = \frac{8}{\pi^2}\sum_{m=1}^{\infty}\frac{1-\kappa_{bi}(m\boldsymbol{X}-m\boldsymbol{Y})}{4m^2-1},$$

where $h(\boldsymbol{X}) \triangleq \frac{1}{2}(1 + sgn(cos(\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v} + b) + t))$, $\boldsymbol{w}, \boldsymbol{v} \sim \mathcal{N}(0, \boldsymbol{I})$, $b \sim Unif[0, 2\pi]$, *and* $t \sim Unif[-1, 1]$.

Proof. Using Lemma 1, we can show that $\mathbb{E}_{\boldsymbol{w},\boldsymbol{v},b,t}\left[\mathbb{I}_{h(\boldsymbol{X})\neq h(\boldsymbol{Y})}\right] = \frac{1}{2}\mathbb{E}_{\boldsymbol{w},\boldsymbol{v},b}|\cos(\boldsymbol{w}^\top\boldsymbol{X}\boldsymbol{v}+b)-\cos(\boldsymbol{w}^\top\boldsymbol{Y}\boldsymbol{v}+b)|$. By using a trigonometry identity,

$$\frac{1}{2}\mathbb{E}_{b,\boldsymbol{w},\boldsymbol{v}}|\cos(\boldsymbol{w}^\top\boldsymbol{X}\boldsymbol{v}+b)-\cos(\boldsymbol{w}^\top\boldsymbol{Y}\boldsymbol{v}+b)| = \frac{2}{\pi}\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}}\left|\sin\left(\frac{\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}}{2}\right)\right|.$$

By [1], we use Fourier series of $g(\tau) = |\sin(\tau)|$:

$$g(\tau) = \frac{4}{\pi}\sum_{m=1}^{\infty}\frac{1-\cos(2m\tau)}{4m^2-1}.$$

This formula leads to the following equation:

$$\mathbb{E}_{\boldsymbol{w},\boldsymbol{v},b,t}\left[\mathbb{I}_{h(\boldsymbol{X})\neq h(\boldsymbol{Y})}\right] = \frac{8}{\pi^2}\sum_{m=1}^{\infty}\frac{1-\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}}\left[\cos(m\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v})\right]}{4m^2-1}$$

According to the proof of Lemma 1 described in the paper, we know that $\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}}\cos(m\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}) = \kappa_{bi}(m\boldsymbol{X}-m\boldsymbol{Y})$, which completes the proof. Q.E.D.

3 Bounds on the Covariance by Bilinear Projections

Lemma 3 *Given a datum as* $\boldsymbol{X} \in \mathcal{R}^{d_1 \times d_2}$ *and bilinear projections,* $\boldsymbol{w}, \boldsymbol{v}_1, \boldsymbol{v}_2$*, are drawn from the* $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$*,* $\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2}\big[cos(m\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_1)cos(n\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_2)\big] = \kappa_{bi}(\sqrt{(m^2+n^2)}\boldsymbol{X})$.

Proof.

$$\begin{aligned}
& \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2}\big[\cos(m\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_1)\cos(n\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_2)\big] \\
= & \int \big[\int \cos(m\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_1)p(\boldsymbol{v}_1)d\boldsymbol{v}_1\big]\big[\int \cos(m\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_2)p(\boldsymbol{v}_2)d\boldsymbol{v}_2\big]p(\boldsymbol{w})d\boldsymbol{w} \\
= & \int \kappa_g(m\boldsymbol{w}^\top \boldsymbol{X})\kappa_g(n\boldsymbol{w}^\top \boldsymbol{X})p(\boldsymbol{w})d\boldsymbol{w} \quad \text{(by Lemma 1 in the paper)} \\
= & \int \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}(\boldsymbol{w}^\top[\boldsymbol{I} + m^2\boldsymbol{X}\boldsymbol{X}^\top + n^2\boldsymbol{X}\boldsymbol{X}^\top]\boldsymbol{w}))d\boldsymbol{w} \\
= & |\boldsymbol{I} + (m^2+n^2)\boldsymbol{X}\boldsymbol{X}^\top|^{-\frac{1}{2}} \triangleq \kappa_{bi}(\sqrt{(m^2+n^2)}\boldsymbol{X}).
\end{aligned}$$

Theorem 1 *(Theorem 3 in the paper) Given the hash functions* $h_1(\cdot)$ *and* $h_2(\cdot)$*, the upper bound on the covariance between the two bits is derived as*

$$cov(\cdot) \le (\frac{64}{\pi^4})\Big[(\sum_{m=1}^{\infty} \frac{\kappa_g(vec(\boldsymbol{X}-\boldsymbol{Y}))^{0.79m^2}}{4m^2-1})^2 - (\sum_{m=1}^{\infty} \frac{\kappa_g(vec(\boldsymbol{X}-\boldsymbol{Y}))^{m^2}}{4m^2-1})^2\Big],$$

where $\kappa_g(\cdot)$ *is the Gaussian kernel and* $cov(\cdot)$ *is the covariance between two bits defined as*

$$\begin{aligned}
cov(\cdot) &= \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2,b_1,b_2,t_1,t_2}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big] \\
&- \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,b_1,t_1}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\big]\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_2,b_2,t_2}\big[\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big], \\
h_1(\boldsymbol{X}) &= sgn\Big(\cos(\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_1 + b_1) + t_1\Big), \\
h_2(\boldsymbol{X}) &= sgn\Big(\cos(\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_2 + b_2) + t_2\Big).
\end{aligned}$$

Proof. First, we want to derive the first term in the covariance in terms of $\kappa_{bi}(\cdot)$.

$$\begin{aligned}
& \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2,b_1,b_2,t_1,t_2}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big] \\
= & \frac{1}{4}\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2,b_1,b_2}\big[|\cos(\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_1 + b_1) - \cos(\boldsymbol{w}^\top \boldsymbol{Y}\boldsymbol{v}_1 + b_1)||\cos(\boldsymbol{w}^\top \boldsymbol{X}\boldsymbol{v}_2 + b_2) - \cos(\boldsymbol{w}^\top \boldsymbol{Y}\boldsymbol{v}_2 + b_2)|\big] \quad (\because \text{Lemma 1}) \\
= & \frac{4}{\pi^2}\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2}\big[|\sin(\frac{\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}_1}{2})\sin(\frac{\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}_2}{2})|\big] \\
= & (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty} \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2}\big[(\frac{1-\cos(m\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}_1)}{4m^2-1})(\frac{1-\cos(n\boldsymbol{w}^\top(\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{v}_2)}{4n^2-1})\big]
\end{aligned}$$

Using the Lemma 2, the first term in the covariance can be represented in terms of $\kappa_{bi}(\cdot)$:

$$\begin{aligned}
& \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,b_1,b_2,t_1,t_2}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big] \\
= & (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty} \frac{1}{4m^2-1}\frac{1}{4n^2-1}\big(1 - \kappa_{bi}(m\boldsymbol{X} - m\boldsymbol{Y}) - \kappa_{bi}(n\boldsymbol{X} - n\boldsymbol{Y}) + \kappa_{bi}(\sqrt{(m^2+n^2)}(\boldsymbol{X}-\boldsymbol{Y}))\big)
\end{aligned}$$

The second term in the covariance is also represented in terms of $\kappa_{bi}(\cdot)$:

$$\begin{aligned}
& \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,b_1,t_1}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\big]\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_2,b_2,t_2}\big[\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big] \\
= & (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty} \frac{1}{4m^2-1}\frac{1}{4n^2-1}\big(1 - \kappa_{bi}(m(\boldsymbol{X}-\boldsymbol{Y})) - \kappa_{bi}(n(\boldsymbol{X}-\boldsymbol{Y})) + \kappa_{bi}(m(\boldsymbol{X}-\boldsymbol{Y}))\kappa_{bi}(n(\boldsymbol{X}-\boldsymbol{Y}))\big)
\end{aligned}$$

Therefore, the covariance between two bits is computed as

$$\begin{aligned}
\text{cov}(\cdot) &= (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty}\frac{1}{4m^2-1}\frac{1}{4n^2-1}\big[\kappa_{bi}(\sqrt{(m^2+n^2)}(\boldsymbol{X}-\boldsymbol{Y}))-\kappa_{bi}(m(\boldsymbol{X}-\boldsymbol{Y}))\kappa_{bi}(n(\boldsymbol{X}-\boldsymbol{Y}))\big] \\
&\leq (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty}\frac{1}{4m^2-1}\frac{1}{4n^2-1}\big[\kappa_g(\text{vec}(\sqrt{(m^2+n^2)}(\boldsymbol{X}-\boldsymbol{Y})))^{0.79}-\kappa_g(\text{vec}(m\boldsymbol{X}-m\boldsymbol{Y}))\kappa_g(\text{vec}(n\boldsymbol{X}-n\boldsymbol{Y}))\big] \\
&= (\frac{64}{\pi^4})\sum_{m,n=1}^{\infty}\frac{1}{4m^2-1}\frac{1}{4n^2-1}\big[\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{0.79(m^2+n^2)}-\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{m^2}\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{n^2}\big] \\
&= (\frac{64}{\pi^4})[(\sum_{m=1}^{\infty}\frac{\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{0.79m^2}}{4m^2-1})^2-(\sum_{m=1}^{\infty}\frac{\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{m^2}}{4m^2-1})^2],
\end{aligned}$$

where the second inequality is given by Lemma 1 in the paper ($\kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y})) \leq \kappa_{bi}(\boldsymbol{X}-\boldsymbol{Y}) \leq \kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{0.79}$) and the third equality is given by $\kappa_g(\text{vec}(m\boldsymbol{X}-m\boldsymbol{Y})) = \kappa_g(\text{vec}(\boldsymbol{X}-\boldsymbol{Y}))^{m^2}$. The lower bound can be derived in a similar way.

Corollary 1 *Given the hash functions $h_1(\cdot)$ and $h_2(\cdot)$, the lower bound on the covariance between the two bits is derived as*

$$cov(\cdot) \geq (\frac{64}{\pi^4})\Big[(\sum_{m=1}^{\infty}\frac{\kappa_g(vec(\boldsymbol{X}-\boldsymbol{Y}))^{m^2}}{4m^2-1})^2-(\sum_{m=1}^{\infty}\frac{\kappa_g(vec(\boldsymbol{X}-\boldsymbol{Y}))^{0.79m^2}}{4m^2-1})^2\Big],$$

where $\kappa_g(\cdot)$ is the Gaussian kernel and $cov(\cdot)$ is the covariance between two bits defined as

$$\begin{aligned}
cov(\cdot) &= \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,\boldsymbol{v}_2,b_1,b_2,t_1,t_2}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big] \\
&- \mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_1,b_1,t_1}\big[\mathbb{I}_{h_1(\boldsymbol{X})\neq h_1(\boldsymbol{Y})}\big]\mathbb{E}_{\boldsymbol{w},\boldsymbol{v}_2,b_2,t_2}\big[\mathbb{I}_{h_2(\boldsymbol{X})\neq h_2(\boldsymbol{Y})}\big], \\
h_1(\boldsymbol{X}) &= sgn\Big(\cos(\boldsymbol{w}^\top\boldsymbol{X}\boldsymbol{v}_1+b_1)+t_1\Big), \\
h_2(\boldsymbol{X}) &= sgn\Big(\cos(\boldsymbol{w}^\top\boldsymbol{X}\boldsymbol{v}_2+b_2)+t_2\Big).
\end{aligned}$$

References

[1] M. Raginsky and S. Lazebnik. Locality-sensitive binary codes from shift-invariant kernels. In *Advances in Neural Information Processing Systems (NIPS)*, volume 22. MIT Press, 2009.