Supplementary Material:
On Symmetric and Asymmetric LSHs for Inner Product Search

Behnam Neyshabur BNEYSHABUR@TTIC.EDU
Nathan Srebro NATI@TTIC.EDU
Toyota Technological Institute at Chicago, Chicago, IL 60637, USA

1. Another variant

To benefit from the empirical advantages of random projection hashing, Shrivastava and Li (2014) also proposed a modified asymmetric LSH, which we refer to here as SIGN-ALSH(SL). SIGN-ALSH(SL) uses two different mappings $P(x)$, $Q(q)$, similar to those of L2-ALSH(SL), but then uses a random projection hash $h_a(x)$, as is the one used by SIMPLE-LSH, instead of the quantized hash used in L2-ALSH(SL). In this appendix we show that our theoretical observations about L2-ALSH(SL) are also valid for SIGN-ALSH(SL).

SIGN-ALSH(SL) uses the pair of mappings:

$$\begin{aligned} P(x) &= [Ux; 1/2 - \|Ux\|^2; \dots; 1/2 - \|Ux\|^{2^m}] \\ Q(y) &= [y; 0; 0; \dots; 0], \end{aligned} \tag{1}$$

where m and U are parameters, as in L2-ALSH(SL). SIGN-ALSH(LS) is then given by $f(x) = h_a(P(x))$, $g(y) = h_a(Q(x))$, where h_a is the random projection hash given in (11). SIGN-ALSH(LS) therefor depends on two parameters, and uses a binary alphabet $\Gamma = \{\pm 1\}$.

In this section, we show that, like L2-ALSH(LS), SIGN-ALSH(LS) is not a universal ALSH over $\mathcal{X}_\bullet, \mathcal{Y}_\circ$, and moreover for any $S > 0$ and $0 < c < 1$ it is not an (S, cS)-ALSH over $\mathcal{X}_\bullet = \mathcal{Y}_\bullet$:

Lemma 1. *For any m, U, r, and for any $0 < S < 1$ and*

$$\min\left\{ \sqrt{1 - \frac{U^{2^{m+1}}(1 - S^{2^{m+1}})}{U^{2^{m+1}} + m/4}}, \frac{\sqrt[2^{m+1}]{\frac{(m/2)}{2^{m+1}-2}}}{SU} \right\} \le c < 1$$

SIGN-ALSH(SL) *is not an (S, cS)-ALSH for inner product similarity over $\mathcal{X}_\bullet = \{x | \|x\| \le 1\}$ and $\mathcal{Y}_\circ = \{q | \|q\| = 1\}$.*

Proof. Assume for contradiction that:

$$\sqrt{1 - \frac{U^{2^{m+1}}(1 - S^{2^{m+1}})}{U^{2^{m+1}} + m/4}} \le c < 1$$

and SIGN-ALSH(SL) is an (S, cS)-ALSH. For any query point $q \in \mathcal{Y}_\circ$, let $x \in \mathcal{X}_\bullet$ be a vector s.t. $q^\top x = S$ and $\|x\|_2 = 1$ and let $y = cSq$, so that $q^\top y = cS$. We have that:

$$\begin{aligned} \frac{(P(y)^\top Q(q))^2}{\|P(y)\|^2} &= \frac{c^2 S^2 U^2}{m/4 + \|y\|^{2^{m+1}}} \\ &= \frac{c^2 S^2 U^2}{m/4 + (cSU)^{2^{m+1}}} \end{aligned}$$

Using $1 - \frac{U^{2^{m+1}}(1 - S^{2^{m+1}})}{U^{2^{m+1}} + m/4} \le c^2 < 1$:

$$\begin{aligned} &> \frac{c^2 S^2 U^2}{m/4 + (SU)^{2^{m+1}}} \\ &\ge \frac{S^2 U^2}{m/4 + U^{2^{m+1}}} \\ &= \frac{(P(x)^\top Q(q))^2}{\|P(x)\|^2} \end{aligned}$$

The monotonicity of $1 - \frac{\cos^{-1}(x)}{\pi}$ establishes a contradiction. To get the other bound on c, let $\alpha_m = \sqrt[2^{m+1}]{\frac{(m/2)}{2^{m+1}-2}}$ and assume for contradiction that:

$$\frac{\alpha_m}{SU} = \frac{\sqrt[2^{m+1}]{\frac{(m/2)}{2^{m+1}-2}}}{SU} \le c < 1$$

and SIGN-ALSH(SL) is an (S, cS)-ALSH. For any query point $q \in \mathcal{Y}_\circ$, let $x \in \mathcal{X}_\bullet$ be a vector s.t. $q^\top x = S$ and $\|x\|_2 = 1$ and let $y = (\alpha_m / U) q$. By the monotonicity of $1 - \frac{\cos^{-1}(x)}{\pi}$, to get a contradiction is enough to show that

$$\frac{P(x)^\top Q(q)}{\|P(x)\|} \le \frac{P(y)^\top Q(q)}{\|P(y)\|}$$

We have:

$$\begin{aligned} \frac{(P(y)^\top Q(q))^2}{\|P(y)\|^2} &= \frac{\alpha_m^2}{m/4 + \|\alpha_m\|^{2^{m+1}}} \\ &= \frac{\sqrt[2^m]{\frac{(m/2)}{2^{m+1}-2}}}{m/4 + \frac{(m/2)}{2^{m+1}-2}} \end{aligned}$$

Since this is the maximum value of the function $f(U) = U^2/(m/4 + U^{2^{m+1}})$:

$$\begin{aligned} &\geq \frac{U^2}{m/4 + U^{2^{m+1}}} \\ &\geq \frac{S^2U^2}{m/4 + U^{2^{m+1}}} \\ &= \frac{(P(x)^\top Q(q))^2}{\|P(x)\|^2} \end{aligned}$$

which is a contradiction. □

Corollary 1.1. *For any U, m and r, SIGN-ALSH(SL) is not a universal ALSH for inner product similarity over $\mathcal{X}_\bullet = \{x \| \|x\| \leq 1\}$ and $\mathcal{Y}_\circ = \{q \| \|q\| = 1\}$. Furthermore, for any $c < 1$, and any choice of U, m, r there exists $0 < S < 1$ for which SIGN-ALSH(SL) is not an (S, cS)-ALSH over $\mathcal{X}_\bullet, \mathcal{Y}_\circ$, and for any $S < 1$ and any choice of U, m, r there exists $0 < c < 1$ for which SIGN-ALSH(SL) is not an (S, cS)-ALSH over $\mathcal{X}_\bullet, \mathcal{Y}_\circ$.*

Lemma 2. *For any $S > 0$ and $0 < c < 1$ there are no U and m such that SIGN-ALSH(SL) is an (S, cS)-ALSH for inner product similarity over $\mathcal{X}_\bullet = \mathcal{Y}_\bullet = \{x \mid \|x\| \leq 1\}$.*

Proof. Similar to the proof of Theorem 5.2, for any $S > 0$ and $0 < c < 1$, let q_1 and x_1 be unit vectors such that $q_1^\top x_1 = S$. Let x_2 be a unit vector and define $q_2 = cSx_2$. For any U and m:

$$\begin{aligned} \frac{P(x_2)^\top Q(q_2)}{\|P(x_2)\| \|Q(q_2)\|} &= \frac{cSU}{cS\sqrt{m/4 + \|U\|^{2^{m+1}}}} \\ &= \frac{U}{\sqrt{m/4 + \|U\|^{2^{m+1}}}} \\ &\geq \frac{SU}{\sqrt{m/4 + \|U\|^{2^{m+1}}}} \\ &= \frac{P(x_1)^\top Q(q_1)}{\|P(x_1)\| \|Q(q_1)\|} \end{aligned}$$

Now, the same arguments as in Lemma 1 using monotonicity of collision probabilities in $\|P(x) - Q(q)\|$ establish SIGN-ALSH(SL) is not an (S, cS)-ALSH. □

2. Max-norm and margin complexity

MAX-NORM

The max-norm (aka $\gamma_2{:}\ell_1{\to}\ell_\infty$ norm) is defined as (Srebro et al., 2005):

$$\|X\|_{\max} = \min_{X=UV^\top} \max(\|U\|_{2,\infty}^2, \|V\|_{2,\infty}^2) \tag{2}$$

where $\|U\|_{2,\infty}$ is the maximum over ℓ_2 norms of rows of matrix U, i.e. $\|U\|_{2,\infty} = \max_i \|U[i]\|$.

For any pair of sets $(\{x_i\}_{1\leq i\leq n}, \{y_i\}_{1\leq i\leq m})$ and hashes (f, g) over them, let P be the collision probability matrix, i.e. $P(i,j) = \mathbb{P}[f(x_i) = g(y_j)]$. In the following lemma we prove that $\|P\|_{\max} \leq 1$:

Lemma 3. *For any two sets of objects and hashes over them, if P is the collision probability matrix, then $\|P\|_{\max} \leq 1$.*

Proof. For each f and g, define the following biclustering matrix:

$$\kappa_{f,g}(i,j) = \begin{cases} 1 & f(x_i) = g(y_j) \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

For any function $f : \mathcal{Z} \to \Gamma$, let $R_f \in \{0,1\}^{n\times|\Gamma|}$ be the indicator of the values of function f:

$$R_h(i,\gamma) = \begin{cases} 1 & h(x_i) = \gamma \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

and define $R_g \in \{0,1\}^{m\times|\Gamma|}$ similarly. It is easy to show that $\kappa_{f,g} = R_f R_g^\top$ and since $\|R_f\|_{2,\infty} = \|R_g\|_{2,\infty} = 1$, by the definition of the max-norm, we can conclude that $\|\kappa_{f,g}\|_{\max} \leq 1$. But the collision probabilities are given by $P = \mathbb{E}[\kappa_{f,g}]$, and so by convexity of the max-norm and Jensen's inequality, $\|P\|_{\max} = \|\mathbb{E}[\kappa_{f,g}]\|_{\max} \leq \mathbb{E}[\|\kappa_{f,g}\|_{\max}] \leq 1$. □

It is also easy to see that $1_{n\times n} = RR^\top$ where $R = 1_{n\times 1}$. Therefore for any $\theta \in \mathbb{R}$,

$$\|\theta_{n\times n}\|_{\max} = \theta \|1_{n\times n}\|_{\max} \leq |\theta|$$

MARGIN COMPLEXITY

For any sign matrix Z, the margin complexity of Z is defined as:

$$\begin{aligned} \min_Y \quad & \|Y\|_{\max} \\ \text{s.t.} \quad & Y(i,j)X(i,j) \geq 1 \quad \forall i,j \end{aligned} \tag{5}$$

Let $Z \in \{\pm 1\}^{N\times N}$ be a sign matrix with +1 on and above the diagonal and -1 below it. Forster et al. (2003) prove that the margin complexity of matrix Z is $\Omega(\log N)$.

References

Forster, J., Schmitt, N., Simon, H., and Suttorp, T. (2003). Estimating the optimal margins of embeddings in euclidean half spaces. *Machine Learning*, 51:263281.

Shrivastava, A. and Li, P. (2014). Improved asymmetric locality sensitive hashing (alsh) for maximum inner product search (mips). *arXiv:1410.5410*.

Srebro, N., Rennie, J., and Jaakkola, T. (2005). Maximum margin matrix factorization. *NIPS*.