Is margin preserved after random projection?

Qinfeng Shi JAVEN.SHI@ADELAIDE.EDU.AU
Chunhua Shen CHUNHUA.SHEN@ADELAIDE.EDU.AU
Rhys Hill RHYS.HILL@ADELAIDE.EDU.AU
Anton van den Hengel ANTON.VANDENHENGEL@ADELAIDE.EDU.AU
Australian Centre for Visual Technologies, The University of Adelaide

Errata

Here we provide corrections and clarifications below:

1. Is acute angle needed to preserve the angle and the inner product?
   Not really! We proved the case for acute angle because a positive margin requires acute angles. However, the result and the proof of Theorem 5 can be extended to the case when $\langle \mathbf{w}, \mathbf{x} \rangle \leq 0$. A simple sanity check is that if $\langle \mathbf{w}, \mathbf{x} \rangle$ (*i.e.* $-\langle \mathbf{w}, -\mathbf{x} \rangle$) is well preserved, then $\langle \mathbf{w}, -\mathbf{x} \rangle$ is well preserved.

   Consequently, our statement in Fig. 2. "Neither angle (c) nor inner product (d) is preserved when the angle is obtuse" and related discussion about "obtuse angle" were incorrect. In fact, when the angle gets close to 90 degree (doesn't matter acute or obtuse angle, positive, or negative inner product), the distortion is big. For an obtuse angle, if it is far away from 90 degree (*i.e.* close to 180 degree or 0), the preservation is good. For example, in Fig. 2 (c) and (d), when $\gamma = -0.062$ the preservation is better than that when $\gamma = -0.0165$.

2. What do you mean by "the normalised margin is more informative than the unnormalised margin"?
   By "more informative", we mean: if you know the normalised margin is big, you know both the normalised margin and unnormalised margin preserve well with high probability after random projection. If you only know the unnormalised margin is big, the unnormalised margin may or may not preserve well (depending on the normalised margin).

3. In Theorem 6, "linearly separable by margin $\gamma - \frac{2\epsilon}{1-\epsilon}$ " should be "linearly separable by margin $(\frac{1+\epsilon}{1-\epsilon})\gamma - \frac{2\epsilon}{1-\epsilon}$ ", although the latter implies the former.

4. Fixing derivation of inequality (9) *i.e.*

$$\| \frac{\mathbf{R}\,\mathbf{x}}{\|\,\mathbf{x}\,\|} - \frac{\mathbf{R}\,\mathbf{w}}{\|\,\mathbf{w}\,\|} \|^2 \leq \| \sqrt{(1+\epsilon)} (\frac{\mathbf{R}\,\mathbf{x}}{\|\,\mathbf{R}\,\mathbf{x}\,\|} - \frac{\mathbf{R}\,\mathbf{w}}{\|\,\mathbf{R}\,\mathbf{w}\,\|}) \|^2 + (\sqrt{(1+\epsilon)} - \sqrt{(1-\epsilon)})^2. \tag{9}$$

   Proof Let $\mathbf{a} = \frac{\mathbf{R}\,\mathbf{x}}{\|\,\mathbf{R}\,\mathbf{x}\,\|}$, $\mathbf{b} = \frac{\mathbf{R}\,\mathbf{w}}{\|\,\mathbf{R}\,\mathbf{w}\,\|}$, $\alpha = \frac{\|\,\mathbf{R}\,\mathbf{x}\,\|}{\|\,\mathbf{x}\,\|}$, $\beta = \frac{\|\,\mathbf{R}\,\mathbf{w}\,\|}{\|\,\mathbf{w}\,\|}$ and $\eta = (\sqrt{(1+\epsilon)} - \sqrt{(1-\epsilon)})^2$. We know $\alpha, \beta \in [\sqrt{1-\epsilon}, \sqrt{1+\epsilon}]$ via inequality (7), thus $\alpha\beta \leq 1+\epsilon$. We also know $\mathbf{a}^2 = 1$, $\mathbf{b}^2 = 1$ and $\mathbf{a}^\intercal\,\mathbf{b} \leq 1$ by definition. RHS of (9) $-$ LHS of (9) yields,

$$\begin{aligned}
&\eta + \| \sqrt{(1+\epsilon)} (\mathbf{a} - \mathbf{b}) \|^2 - \|\,\mathbf{a}\,\alpha - \mathbf{b}\,\beta \|^2 \\
&= \eta + (1+\epsilon)(\mathbf{a}^2 + \mathbf{b}^2) - (1+\epsilon)(2\,\mathbf{a}^\intercal\,\mathbf{b}) - \alpha^2\,\mathbf{a}^2 - \beta^2\,\mathbf{b}^2 + 2\alpha\beta\,\mathbf{a}^\intercal\,\mathbf{b} \\
&= \eta + 2(1+\epsilon) - 2\,\mathbf{a}^\intercal\,\mathbf{b}[(1+\epsilon) - \alpha\beta] - \alpha^2 - \beta^2 \\
&\geq \eta + 2(1+\epsilon) - 2[(1+\epsilon) - \alpha\beta] - \alpha^2 - \beta^2 \qquad (\because [(1+\epsilon) - \alpha\beta] > 0, \mathbf{a}^\intercal\,\mathbf{b} \leq 1) \\
&= \eta - (\alpha - \beta)^2 \geq 0.
\end{aligned}$$

   Thus inequality (9) holds. This proof is suggested by Lijun Zhang (MSU).