Unifying Cosine and PLDA Back-ends for Speaker Verification: Appendix for proofs

Zhiyuan PENG, jerrypeng@1937@gmail.com

April 20, 2022

Q1: Consider a set of N embeddings $\mathcal{X} = \{x_n\}_{n=1}^{n_1}$ that come from the same speaker y. Each embedding x_n is of D dimensions. Prove that

$$\log p(\mathcal{X}) = \frac{1}{2}\left(n_1^2 \mu_1^T W (B + n_1 W)^{-1} W \mu_1 - \sum_{n=1}^{n_1} x_n^T W x_n + C\right) \tag{1}$$

$$C = \log|B| + n_1 \log|W| - \log|B + n_1 W| - n_1 D \log(2\pi) \tag{2}$$

where $\mu_1 = \frac{1}{n_1}\sum_{n=1}^{n_1} x_n$.

Proof: Recall Eq.(3) in the original paper (note that $\mu = 0$ as mentioned in Section 2.3),

$$\begin{aligned} p(\mathcal{X}, y) &\propto \exp\left(-\frac{1}{2}(y^T B y + \sum_{n=1}^{n_1}(x_n - y)^T W (x_n - y)\right) \\ &\propto \exp\left(-\frac{1}{2}\left[y^T(B + n_1 W)y - 2\sum_{n=1}^{n_1} x_n^T W y + \sum_{n=1}^{n_1} x_n^T W x_n\right]\right) \end{aligned} \tag{3}$$

Let $\mu_* = n_1(B + n_1 W)^{-1} W \mu_1$, it can be derived that,

$$p(\mathcal{X}, y) \propto \exp\left(-\frac{1}{2}\left[(y - \mu_*)^T (B + n_1 W)(y - \mu_*) - \mu_*^T (B + n_1 W)\mu_* + \sum_{n=1}^{n_1} x_n^T W x_n\right]\right) \tag{4}$$

Thus the marginal distribution $p(\mathcal{X})$ can be expressed as,

$$p(\mathcal{X}) \propto \exp\left(-\frac{1}{2}\left[\sum_{n=1}^{n_1} x_n^T W x_n - \mu_*^T (B + n_1 W)\mu_*\right]\right) \tag{5}$$

Taking logarithm to both sides, it can be shown that,

$$\log p(\mathcal{X}) = \frac{1}{2}\left(\mu_*^T (B + n_1 W)\mu_* - \sum_{n=1}^{n_1} x_n^T W x_n + C_{\mathcal{X}}\right) \tag{6}$$

where $C_{\mathcal{X}}$ is a constant. The first two terms in the RHS of (6) exactly match those in the RHS of (1). Next we prove that $C_{\mathcal{X}} = C$.

Consider a variable z that is multivariate Gaussian,

$$p(z) = \frac{1}{(2\pi)^{D/2}|G|^{-1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T G(x-\mu)\right) \tag{7}$$

Taking logarithm to both sides, it can be shown that,

$$\log p(z) = \frac{1}{2}\left(\log|G| - D\log(2\pi) - (x-\mu)^T G(x-\mu)\right) \tag{8}$$

Let C_z be the double of the constant part of $\log p(z)$, e.g., $C_z = \log|G| - D\log(2\pi)$.

According to the bayes' theorem,

$$\log p(\mathcal{X}) = \log p(\mathcal{X}, y) - \log p(y|\mathcal{X}) \tag{9}$$

Thus, it can be derived that,

$$C_{\mathcal{X}} = C_{\mathcal{X},y} - C_{y|\mathcal{X}} \tag{10}$$

According to Eq.(3),

$$\begin{aligned} C_{\mathcal{X},y} &= C_y + C_{\mathcal{X}|y} \\ &= (\log|B| + n_1 \log|W| - (n_1+1)D\log(2\pi)) \end{aligned} \tag{11}$$

and,

$$\begin{aligned} &p(y|\mathcal{X}) \propto \exp\left(-\frac{1}{2}\left[y^T(B+n_1W)y - 2\sum_{n=1}^{n_1} x_n^T W y\right]\right) \\ &\Rightarrow C_{y|\mathcal{X}} = \log|B+n_1W| - D\log(2\pi) \end{aligned} \tag{12}$$

Thus,

$$C_{\mathcal{X}} = \log|B| + n_1\log|W| - \log|B+n_1W| - n_1 D\log(2\pi) \tag{13}$$

This completes the proof.

Q2: Given two embeddings x_i, x_j, show that the LLR of PLDA can be expressed as

$$S_{\text{PLDA}}(x_i, x_j) \doteq \frac{1}{2}\left(x_i^T Q x_i + x_j^T Q x_j + 2x_i^T P x_j\right) \tag{14}$$

, where $\doteq$ means equivalence up to a negligible additive constant, and

$$Q = W((B + 2W)^{-1} - (B + W)^{-1})W \tag{15}$$

$$P = W(B + 2W)^{-1}W \tag{16}$$

Proof: According to Eq.(4) in the original paper,

$$S_{\text{PLDA}}(x_i, x_j) = \log \frac{p(x_i, x_j)}{p(x_i)p(x_j)} \tag{17}$$

With Eq.(1), it is known that,

$$\begin{aligned}
\log p(x_i, x_j) &\doteq \frac{1}{2}\left((x_i + x_j)^T W(B + 2W)^{-1}W(x_i + x_j) - x_i^T W x_i - x_j^T W x_j\right) \\
\log p(x_i) &\doteq \frac{1}{2}\left(x_i^T W(B + W)^{-1}W x_i - x_i^T W x_i\right) \\
\log p(x_j) &\doteq \frac{1}{2}\left(x_j^T W(B + W)^{-1}W x_j - x_j^T W x_j\right)
\end{aligned} \tag{18}$$

Putting Eq.(18) into Eq.(17), Eq.(14-16) can be derived.

Q3: Consider two sets of embeddings $\mathcal{X}_1$ and $\mathcal{X}_2$ of size K_1 and K_2, respectively. Their centroids are denoted by μ_1 and μ_2. Show that

$$S_{\text{PLDA}}(\mathcal{X}_1, \mathcal{X}_2) = \frac{K_1K_2}{1 + K_1 + K_2} S_{\cos}(\mu_1, \mu_2) + \frac{1}{2}C(K_1, K_2) \tag{19}$$

$$\begin{aligned}
C(K_1, K_2) = &\frac{K_1^2 + K_2^2}{1 + K_1 + K_2} - \frac{K_1^2}{1 + K_1} - \frac{K_2^2}{1 + K_2} \\
&+ \log(1 + \frac{K_1K_2}{1 + K_1 + K_2})
\end{aligned} \tag{20}$$

under the condition of $W = B = I$.

Proof: The LLR of PLDA for two sets of embeddings $\mathcal{X}_1, \mathcal{X}_2$ is defined as:

$$S_{\text{PLDA}}(\mathcal{X}_1, \mathcal{X}_2) = \log \frac{p(\mathcal{X}_1, \mathcal{X}_2)}{p(\mathcal{X}_1)p(\mathcal{X}_2)} \tag{21}$$

Given the condition $W = B = I$, according to Eq.(6),

$$\begin{aligned}
\log p(\mathcal{X}_1) &= \frac{1}{2}\left(\frac{K_1^2}{K_1 + 1}\mu_1^T\mu_1 - \sum_{n=1}^{K_1} {x_n^1}^T x_n^1 + C_{\mathcal{X}_1}\right) \\
\log p(\mathcal{X}_2) &= \frac{1}{2}\left(\frac{K_2^2}{K_2 + 1}\mu_2^T\mu_2 - \sum_{n=1}^{K_2} {x_n^2}^T x_n^2 + C_{\mathcal{X}_2}\right) \\
\log p(\mathcal{X}_1, \mathcal{X}_2) &= \frac{1}{2}\left(\frac{(K_1 + K_2)^2}{K_1 + K_2 + 1}\mu_{1,2}^T\mu_{1,2} - \sum_{n=1}^{K_1} {x_n^1}^T x_n^1 - \sum_{n=1}^{K_2} {x_n^2}^T x_n^2 + C_{\mathcal{X}_1 + \mathcal{X}_2}\right)
\end{aligned} \tag{22}$$

Putting Eq.(22) into Eq.(21), it can be derived that,

$$S_{\mathrm{PLDA}}(\mathcal{X}_1, \mathcal{X}_2) = \frac{1}{2}\left(\frac{(K_1+K_2)^2}{K_1+K_2+1}\mu_{1,2}^T\mu_{1,2} - \frac{K_1^2}{K_1+1}\mu_1^T\mu_1 - \frac{K_2^2}{K_2+1}\mu_2^T\mu_2 + C_*(K_1,K_2)\right) \tag{23}$$

, where $C_*(K_1,K_2) = C_{\mathcal{X}_1+\mathcal{X}_2} - C_{\mathcal{X}_1} - C_{\mathcal{X}_2} = \log(1+\frac{K_1K_2}{1+K_1+K_2})$ and $\mu_{1,2} = \frac{K_1\mu_1+K_2\mu_2}{K_1+K_2}$.
The Eq.(23) can be further simplified as,

$$\begin{aligned} S_{\mathrm{PLDA}}(\mathcal{X}_1, \mathcal{X}_2) =& \frac{K_1K_2}{K_1+K_2+1}\mu_1^T\mu_2 + \frac{1}{2}\left(\frac{K_1^2}{K_1+K_2+1} - \frac{K_1^2}{K_1+1}\right)\mu_1^T\mu_1 \\ &+ \frac{1}{2}\left(\frac{K_2^2}{K_1+K_2+1} - \frac{K_2^2}{K_2+1}\right)\mu_2^T\mu_2 + \frac{1}{2}\log(1+\frac{K_1K_2}{1+K_1+K_2}) \end{aligned} \tag{24}$$

Consider the common case that in PLDA scoring, the centroids are also length-normalized, e.g., $\mu_1^T\mu_1 = 1$. In this regard, the Eq.(24) can be finally simplified as,

$$\begin{aligned} S_{\mathrm{PLDA}}(\mathcal{X}_1, \mathcal{X}_2) &= \frac{K_1K_2}{K_1+K_2+1}\mu_1^T\mu_2 \\ &+ \frac{1}{2}\left(\frac{K_1^2+K_2^2}{K_1+K_2+1} - \frac{K_1^2}{K_1+1} - \frac{K_2^2}{K_2+1} + \log(1+\frac{K_1K_2}{1+K_1+K_2})\right) \\ &= \frac{K_1K_2}{K_1+K_2+1}S_{\cos}(\mu_1,\mu_2) + \frac{1}{2}C(K_1,K_2) \end{aligned} \tag{25}$$