Learnable Adaptive Cosine Estimator (LACE) for Image Classification- Supplementary

Joshua Peeples Connor H. McCurley Sarah Walker Dylan Stewart Alina Zare
Department of Electrical & Computer Engineering, University of Florida
{jpeeples, cmccurley, sarah.walker, d.stewart, azare} @ufl.edu

Appendix

We provide the derivations for the derivatives of the parameters for LACE as discussed in Section 3.1. The derivatives for the target signatures is shown in Section A and the background statistics (*i.e.*, mean and covariance) are shown in Sections B and C.

A. Target Signatures

$$L_{LACE} = \frac{1}{B}\sum_{n=1}^{B} -\log\left(\frac{\exp \hat{\hat{\boldsymbol{s}}}_c^T \hat{\hat{\boldsymbol{x}}}_n}{\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\hat{\boldsymbol{x}}}_n}\right) \tag{1}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T \hat{\hat{\boldsymbol{x}}}_n - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\hat{\boldsymbol{x}}}_n\right)\right) \tag{2}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{x}}}_n{}^T \hat{\hat{\boldsymbol{s}}}_c - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{x}}}_n{}^T \hat{\hat{\boldsymbol{s}}}_j\right)\right) \tag{3}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_c}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_c\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_j}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_j\|}\right)\right)\right) \tag{4}$$

$$\boldsymbol{z}_c = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_c \tag{5}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_c}{\|\boldsymbol{z}_c\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_j}{\|\boldsymbol{z}_j\|}\right)\right)\right) \tag{6}$$

$$\boldsymbol{z}_c' = \frac{\partial}{\partial \boldsymbol{s}_c}\left(\frac{\boldsymbol{z}_c}{\|\boldsymbol{z}_c\|}\right) = \frac{\partial}{\partial \boldsymbol{s}_c}\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_c}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T\boldsymbol{s}_c\|}\right) = \left(\frac{\|\boldsymbol{z}_c\|(\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T) - \frac{\boldsymbol{z}_c\left(\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^T\boldsymbol{s}_c\right)^T}{\|\boldsymbol{z}_c\|}}{\|\boldsymbol{z}_c\|^2}\right) \tag{7}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{s}_c} = \frac{-1}{B}\sum_{n=1}^{B}\frac{\partial}{\partial \boldsymbol{s}_c}\left(\hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_c}{\|\boldsymbol{z}_c\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_j}{\|\boldsymbol{z}_j\|}\right)\right)\right) \tag{8}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{s}_c} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{x}}}_n{}^T\boldsymbol{z}_c' - \frac{\exp \hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_c}{\|\boldsymbol{z}_c\|}\right)\hat{\hat{\boldsymbol{x}}}_n{}^T\boldsymbol{z}_c'}{\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{x}}}_n{}^T\left(\frac{\boldsymbol{z}_j}{\|\boldsymbol{z}_j\|}\right)}\right) \tag{9}$$

B. Background Mean

$$L_{LACE} = \frac{1}{B}\sum_{n=1}^{B} -\log\left(\frac{\exp \hat{\hat{\boldsymbol{s}}}_c^T \hat{\boldsymbol{x}}_n}{\sum_{j=1}^{C} \exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\boldsymbol{x}}_n}\right) \tag{1}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T \hat{\boldsymbol{x}}_n - \log\left(\sum_{j=1}^{C} \exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\boldsymbol{x}}_n\right)\right) \tag{2}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)\|}\right)\right)\right) \tag{3}$$

$$\boldsymbol{m}_b = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b) \tag{4}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right)\right)\right) \tag{5}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\mu}_b} = \frac{-1}{B}\sum_{n=1}^{B}\frac{\partial}{\partial \boldsymbol{\mu}_b}\left(\hat{\hat{\boldsymbol{s}}}_c^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right) - \log\left(\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right)\right)\right) \tag{6}$$

$$\boldsymbol{m}_b' = \frac{\partial}{\partial \boldsymbol{\mu}_b}\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right) = \frac{\partial}{\partial \boldsymbol{\mu}_b}\left(\frac{\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)}{\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b)\|}\right) = \frac{-\|\boldsymbol{m}_b\|\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T - \frac{\boldsymbol{m}_b\left(\boldsymbol{x}_n^T\Sigma_b^{-1} + \Sigma_b^{-1}\boldsymbol{\mu}_b\right)}{\|\boldsymbol{m}_b\|}}{\|\boldsymbol{m}_b\|^2} \tag{7}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\mu}_b} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T\boldsymbol{m}_b' - \frac{\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right)\hat{\hat{\boldsymbol{s}}}_j^T\boldsymbol{m}_b'}{\sum_{j=1}^{C}\exp \hat{\hat{\boldsymbol{s}}}_j^T\left(\frac{\boldsymbol{m}_b}{\|\boldsymbol{m}_b\|}\right)}\right) \tag{8}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\mu}_b} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T\boldsymbol{m}_b' - \sum_{j=1}^{C}\hat{\hat{\boldsymbol{s}}}_j^T\boldsymbol{m}_b'\right) \tag{9}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\mu}_b} = \frac{1}{B}\sum_{n=1}^{B}\sum_{\substack{j=1\\ j\neq c}}^{C}\hat{\hat{\boldsymbol{s}}}_j^T\boldsymbol{m}_b' \tag{10}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\mu}_b} = \frac{1}{B}\sum_{n=1}^{B}\sum_{\substack{j=1\\ j\neq c}}^{C}\hat{\hat{\boldsymbol{s}}}_j^T\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{U}^T(\boldsymbol{x}_n-\boldsymbol{\mu}_b) \tag{11}$$

C. Inverse Background Covariance

Note: In the code, we do not compute the inverse. Instead, by enforcing this metric, we expect the algorithm to learn a matrix that is the inverse background covariance. We take the derivative here with respect to Σ_b^{-1}.

$$L_{LACE} = \frac{1}{B}\sum_{n=1}^{B} -\log\left(\frac{\exp \hat{\hat{\boldsymbol{s}}}_c^T \hat{\boldsymbol{x}}_n}{\sum_{j=1}^{C} \exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\boldsymbol{x}}_n}\right) \tag{1}$$

$$L_{LACE} = \frac{-1}{B}\sum_{n=1}^{B}\left(\hat{\hat{\boldsymbol{s}}}_c^T \hat{\boldsymbol{x}}_n - \log\left(\sum_{j=1}^{C} \exp \hat{\hat{\boldsymbol{s}}}_j^T \hat{\boldsymbol{x}}_n\right)\right) \tag{2}$$

$$\hat{\hat{\boldsymbol{s}}}_c^T \hat{\hat{\boldsymbol{x}}}_n = \frac{\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} \tag{3}$$

$$L_{LACE} = \frac{-1}{B} \sum_{n=1}^{B} \left(\frac{\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} - \log \left(\sum_{j=1}^{C} \exp \frac{\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} \right) \right) \tag{4}$$

$$\mathcal{V}_c' = \frac{\partial}{\partial \boldsymbol{\Sigma}_b^{-1}} \|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\| = \frac{\boldsymbol{s}_c \boldsymbol{s}_c^T \left(\boldsymbol{x}_n - \boldsymbol{\mu}_b)^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b\right) + \boldsymbol{s}_c^T \Sigma_b^{-1} \boldsymbol{s}_c (\boldsymbol{x}_n - \boldsymbol{\mu}_b)(\boldsymbol{x}_n - \boldsymbol{\mu}_b)^T}{2\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} \tag{5}$$

$$V_c' = \frac{\partial}{\partial \boldsymbol{\Sigma}_b^{-1}} \left(\frac{\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} \right) = \left(\frac{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\| (\boldsymbol{x}_n - \boldsymbol{\mu}_b) \boldsymbol{s}_c^T - \boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b) \mathcal{V}_c'}{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|^2} \right) \tag{6}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\Sigma}_b^{-1}} = \frac{-1}{B} \sum_{n=1}^{B} \frac{\partial}{\partial \boldsymbol{\Sigma}_b^{-1}} \left(\frac{\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_c^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} - \log \left(\sum_{j=1}^{C} \exp \frac{\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|} \right) \right) \tag{7}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\Sigma}_b^{-1}} = \frac{-1}{B} \sum_{n=1}^{B} \left(V_c' - \frac{\sum_{j=1}^{C} \exp \frac{\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|}}{\sum_{j=1}^{C} \exp \frac{\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)}{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|}} V_j' \right) \tag{8}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\Sigma}_b^{-1}} = \frac{-1}{B} \sum_{n=1}^{B} \left(V_c' - \sum_{j=1}^{C} V_j' \right) \tag{9}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\Sigma}_b^{-1}} = \frac{1}{B} \sum_{n=1}^{B} \sum_{\substack{j=1 \\ j \neq c}}^{C} V_j' \tag{10}$$

$$\frac{\partial L_{LACE}}{\partial \boldsymbol{\Sigma}_b^{-1}} = \frac{1}{B} \sum_{n=1}^{B} \sum_{\substack{j=1 \\ j \neq c}}^{C} \left(\frac{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\| (\boldsymbol{x}_n - \boldsymbol{\mu}_b) \boldsymbol{s}_j^T - \boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b) \mathcal{V}_j'}{\|\boldsymbol{s}_j^T \Sigma_b^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}_b)\|^2} \right) \tag{11}$$