Supplemental Materials for Affine-Constrained Group Sparse Coding and Its Application to Image-Based Classifications

Yu-Tseh Chi* Mohsen Ali* Muhammad Rushdi† Jeffrey Ho

Department of Computer and Information Science and Engineering, University of Florida

†Department of Biomedical and Systems Engineering, Cairo University, Giza, Egypt

ychi@cise.ufl.edu moali@cise.ufl.edu mrushdi@eng.cu.edu.eg jho@cise.ufl.edu

Abstract

Supplementary materials for the ICCV paper titled Affine-Constrained Group Sparse Coding and Its Application to Image-Based Classifications.

1. Optimization Details

The optimization of the convex objective function

$$\mathcal{E}(\mathbf{a}, \mathbf{c}; \mathbf{X}, \mathbf{D}) = \|\mathbf{X}\,\mathbf{a} - \mathbf{D}\,\mathbf{c}\|^2 + \lambda\,\psi(\mathbf{c}) \tag{1}$$

subject to the nonnegative affine constraint

$$\sum_{i=1}^{k} \mathbf{a}_i = 1,\ \mathbf{a}_1, ..., \mathbf{a}_k \geq 0, \tag{2}$$

is straightforward. The regularizer $\psi(\mathbf{c})$ is a block-based ℓ_1/ℓ_2-norm. We note that "$\mathbf{a}$"-part of the objective function gives a quadratic program with linear constraints (equality and inequalities), while the $\mathbf{c}$-part of the objective function is the same as the typical convex program solved in sparse coding [3]. We use block-coordinate descend with two coordinate-blocks $\mathbf{a}$ and $\mathbf{c}$. The affine constraint is on $\mathbf{a}$ and after each descent step on $\mathbf{a}$, we project the current $\mathbf{a}$ onto the simplex $\sum_{i=1}^{k} \mathbf{a}_i = 1, a_1, ..., a_k \geq 0$. With appropriate step size, this block-coordinate descent algorithm is guaranteed to converge [1].

Projection onto Simplex The only part of the optimization that needs elaboration is the nonnegative affine constraints $\sum_{i=1}^{k} \mathbf{a}_i = 1, \mathbf{a}_1, ...\mathbf{a}_k \geq 0$. For this, we follow the recent work on simplex projection presented in [2]. Let $\mathbf{a} \in \mathbb{R}^n$ and define its projection $\hat{\mathbf{a}}$ on the simplex Δ^n as

$$\hat{\mathbf{a}} = \arg\min_{x \in \Delta^n} \|\mathbf{a} - x\|^2.$$

The main result in [2] shows that this projection can be computed quite easily using a shrinkage-like operation,

$$\hat{\mathbf{a}} = (\mathbf{a} - \mathbf{b})^+,$$

where $(\mathbf{a} - \mathbf{b})^+$ denotes the positive part of $\mathbf{a} - \mathbf{b}$, and the components of the vector $\mathbf{b} \in \mathbb{R}^n$ is determined uniquely by the components of $\mathbf{a}$ (See Theorem 2.2 in [2]).

References

[1] D. Bertsekas. *Convex Analysis and Optimization*. Athena Scientific, 2003.

[2] Y. Chen and X. Ye. *Projection Onto A Simplex*. http://arxiv.org/abs/1101.6081, 2011.

[3] M. Elad. *Sparse and Redundant Representations.* Springer Verlag, 2010.

*These two authors have equal contribution to this paper.