Convolutional K-SVD

Arthur Szlam, Koray Kavukcuoglu, and Yann LeCun
NYU,
715 Broadway,
New York, NY,
aszlam@courant.nyu.edu, koray@courant.nyu.edu,
yann@courant.nyu.edu

February 18, 2010

The K-SVD [1] algorithm is a greedy method for solving the dictionary learning problem

$$arg\min_{W,Z} ||X - WZ||^2_{FRO},$$

such that each column Z_j of Z satisifies $|Z_j|_0 \leq q$. Here X is a $d \times N$ matrix of data written as columns, the dictionary W is a $d \times K$ matrix, and Z is a $K \times N$ matrix of coefficients of the data point in the dictionary W, q is a sparsity parameter, and the error is measured in the Frobenious norm, i.e. the sum of all the entries of the residual squared. The algorithm progresses by alternating updates of the Z, usually via an OMP, and then updating an element of W_i of W with the X corresponding to the nonzero entries in the ith row of Z. We demonstrate that OMP and the K-SVD algorithm translate nicely to the convolutional setting; that is, to solve the problem

$$arg\min_{f,g} ||X - \sum_{j=1}^{K} f_j * g_j||^2_{FRO},$$

$$|\overline{g}|_0 \leq q,$$

similiar to the group l_1 regularized problem considered in [2]. Here $|\overline{g}|_0$ is the total number of nonzero elements in the "cube" of coefficients g. When trained on images, the convolutional structure of the dictionary saves it from having to spend energy on translations, and the resulting model provides a compact representation of edge sets of images.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, 2006.

[2] Koray Kavukcuoglu, Marc'Aurelio Ranzato, Rob Fergus, and Yann LeCun. Learning invariant features through topographic filter maps. In

Proc. International Conference on Computer Vision and Pattern Recognition (CVPR'09). IEEE, 2009.