Convolutional Sparse Coding

A signal $X \in \mathbb{R}^n$ is represented as the sum of small support filters $d_i \in \mathbb{R}^m$, convolved with sparse feature-maps $Z_i \in \mathbb{R}^n$:

$$X = \sum_{i=1}^{m} d_i \ast Z_i.$$  

- Rewrite the CSC in matrix form ($C$ are banded circulant):

$$X = \sum_{i=1}^{m} C \Gamma^i = D \Gamma.$$  

- By permuting the columns of $D$ we obtain the global dictionary:

$$D = \{\Gamma \} = \min_{\beta \Gamma} \frac{1}{2} \|X - D \Gamma\|^2 + \lambda \|\Gamma\|_1$$  

(CSC)  

- CSC poses two main problems – Pursuit ($P_1$) and Dictionary Learning ($D_1$):

$$\hat{\Gamma} = \arg \min_{\Gamma} \frac{1}{2} \|X - D \Gamma\|^2 + \lambda \|\Gamma\|_1 \quad (P_1)$$  

$$\min_{D, \Gamma} \frac{1}{2} \|X - D \Gamma\|^2 + \lambda \|\Gamma\|_1 \quad (D_1)$$  

- Many algorithms were proposed for solving these CSC problems [2,3,4,5]. Most algorithms employ ADMM in the Fourier domain.

Few Definitions

- Stochastic-LoBCoD ($D_1$):

  - The gradient of the CSC objective w.r.t $D_1$:

  $$\nabla D_1 = \sum_{i=1}^{m} p_i \nabla d_i \ast Z_i.$$  

- The gradient $F_{D_1}$ is separable w.r.t the patches and their corresponding filters $d_i$:

  $$D_1 = \text{Proj}(D_1 - \eta \nabla D_1).$$

LoBCoD – Pursuit ($P_1$)

- Proposes Local strategy – split the global sparse vector $\Gamma$ into needles $\alpha_i$:

  $$\hat{\alpha}_i = \arg \min_{\alpha_i} \frac{1}{2} \|X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i \|^2 + \lambda \|\alpha_i\|_1$$  

- Write ($P_1$) in terms of $D_1$ and the needles $\alpha_i$:

  $$\hat{\alpha}_i = \arg \min_{\alpha_i} \frac{1}{2} \|X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i \|^2 + \lambda \|\alpha_i\|_1$$

- Split to $N$ minimization sub-problems:

  $$\alpha_i = \arg \min_{\alpha_i} \frac{1}{2} \|X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i \|^2 + \lambda \|\alpha_i\|_1$$

- The above is equivalent to (local pursuit):

  $$\alpha_i = \arg \min_{\alpha_i} \frac{1}{2} \|X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i \|^2 + \lambda \|\alpha_i\|_1$$

- Every local pursuit step is followed by an update of the residual $R_i(\alpha)$:

  $$L_i = \min_{\alpha_i} \frac{1}{2} \|P_i R_i - D_1 \alpha_i\|^2 + \lambda \|\alpha_i\|_1$$

- Needles with no overlaps are updated in parallel:

  $$L_i = \min_{\alpha_i} \frac{1}{2} \|P_i R_i - D_1 \alpha_i\|^2 + \lambda \|\alpha_i\|_1$$

LoBCoD Algorithm

- For $t = 1:\infty$:

  I. Residual Computation:

  $$R_i \leftarrow R + \sum_{j=1}^{m} p_i^T \nabla d_i \ast Z_i.$$  

  II. Local Sparse Pursuit:

  $$\alpha_i \leftarrow \arg \min_{\alpha_i} \frac{1}{2} \|X - \sum_{j=1}^{n} p_j^T D_i \alpha_i \|^2 + \lambda \|\alpha_i\|_1$$

  III. Reconstructed signal update:

  $$X \leftarrow X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i.$$  

  IV. Residual Signal Update:

  $$R \leftarrow X - \sum_{j=1}^{n} p_j^T D_1 \alpha_i$$

  V. Computation of the gradient:

  $$\nabla d_i \leftarrow \sum_{j=1}^{n} p_j^T \nabla d_i \ast Z_i.$$  

  VI. Dictionary Update:

  $$D_1 \leftarrow \text{Proj}(D_1 - \eta \nabla D_1).$$