Sparse Estimation with Structured Dictionaries

David Wipf
Microsoft Research Asia

Sparse estimation problem:

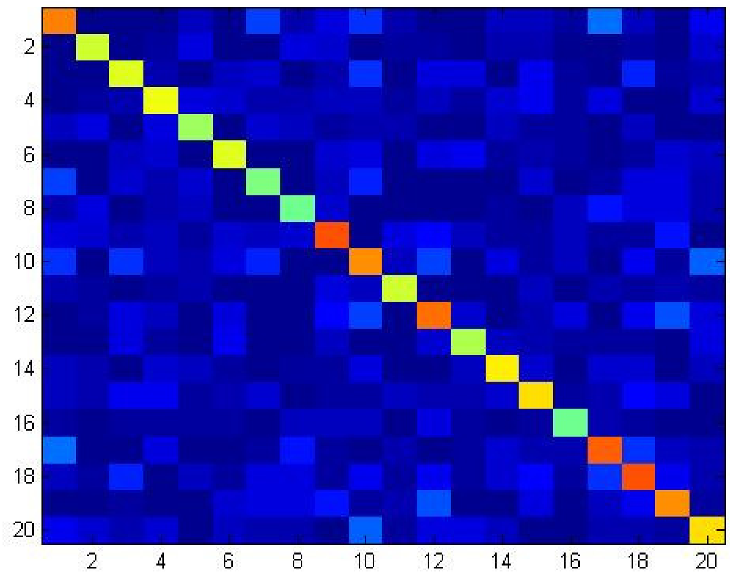
$$\min_{\mathbf{x}} \quad \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \underbrace{\Phi}_{\text{overcomplete dictionary of basis vectors}}\mathbf{x}$$

- Non-convex, combinatorial problem in general.
- Convex relaxation using the ℓ_1 norm produces an equivalent solution if Φ is sufficiently *unstructured*.

Dictionary Correlation Structure

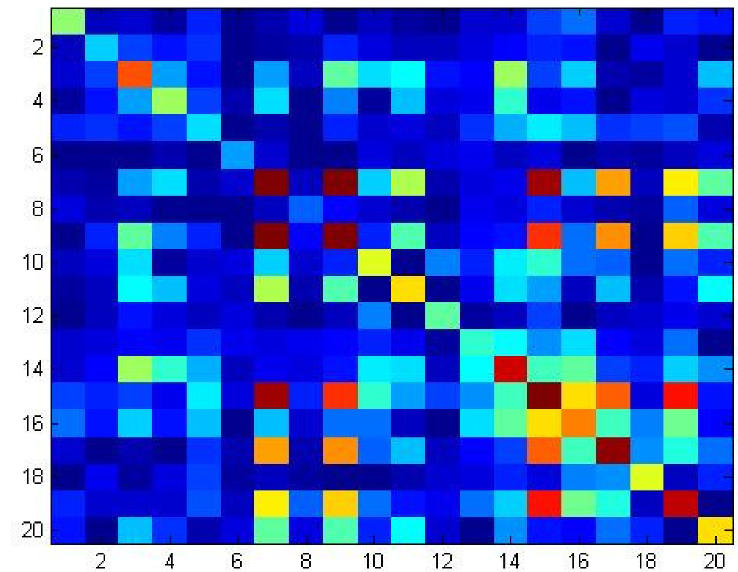
Unstructured

$$\Phi^T \Phi$$



Structured

$$\Phi^T \Phi$$



Examples:

$\Phi_{(unstr)} \sim$ iid $N(0,1)$ entries

$\Phi_{(unstr)} \sim$ random rows of DFT

Example:

$$\Phi_{(str)} = \underbrace{\mathrm{W}}_{\text{arbitrary}} \cdot \Phi_{(unstr)} \cdot \underbrace{\mathrm{D}}_{\text{block diagonal}}$$

New Strategy

- Apply a Φ-dependent projection that maps $\mathbf{x}$ to a new space

$$\mathbf{z} = P_{\Phi}(\mathbf{x})$$

- Use a standard sparsity penalty g in this new space and solve:

$$\min_{\mathbf{x}} \sum_i g(z_i) \quad \text{s.t.} \quad \begin{array}{l} \mathbf{y} = \Phi \mathbf{x} \\ \mathbf{z} = P_{\Phi}(\mathbf{x}) \end{array}$$

The projection operator:

1. Must compensate for dictionary structure.
2. Preserve sparsity, meaning if $\mathbf{z}$ is maximally sparse, $\mathbf{x}$ is also maximally sparse.

Analysis

- Convenient optimization via reweighted ℓ_1 minimization
- Provable performance improvement in certain situations

- **Toy Example:**
- Generate 50-by-100 dictionaries:

$$\Phi_{(\mathrm{unstr})} \sim \mathrm{N}(0,1), \quad \Phi_{(\mathrm{str})} = \Phi_{(\mathrm{unstr})} \cdot \mathrm{D}$$

- Generate a sparse $\mathbf{x}$
- Estimate $\mathbf{x}$ from observations

$$\mathbf{y}_{(\mathrm{unstr})} = \Phi_{(\mathrm{unstr})} \cdot \mathbf{x}\ , \qquad \mathbf{y}_{(\mathrm{str})} = \Phi_{(\mathrm{str})} \cdot \mathbf{x}$$