



Learning computationally efficient dictionaries and their implementation as fast transforms

Luc Le Magoarou, Rémi Gribonval

► **To cite this version:**

Luc Le Magoarou, Rémi Gribonval. Learning computationally efficient dictionaries and their implementation as fast transforms. 2015. <hal-01010577v3>

HAL Id: hal-01010577

<https://hal.inria.fr/hal-01010577v3>

Submitted on 26 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Supplementary material

A. Anonymous

1 Projection operator proof

We want to find the projection operator onto the following set:

$$\mathcal{E} := \{\mathbf{A} \in \mathbb{R}^{n \times n} : \|\mathbf{A}\|_0 \leq p, \|\mathbf{A}\|_F = 1\}, \quad (1)$$

with $p \in \mathbb{N}^*$. We are interested in the projection of some matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ onto the set \mathcal{E} , namely we want to find \mathbf{S}^* such that:

$$\mathbf{S}^* = P_{\mathcal{E}}(\mathbf{S}) \in \arg \min_{\mathbf{U}} \{\|\mathbf{U} - \mathbf{S}\|_F^2 : \mathbf{U} \in \mathcal{E}\} \quad (2)$$

Proposition 1.1. *Projection operator formula.*

$$P_{\mathcal{E}}(\mathbf{S}) = B_1(T_p(\mathbf{S}))$$

where

$$T_p(\mathbf{S}) := \arg \min_{\mathbf{V}} \{\|\mathbf{V} - \mathbf{S}\|_F^2 : \|\mathbf{V}\|_0 \leq p\}$$

and

$$B_1(\mathbf{R}) := \arg \min_{\mathbf{W}} \{\|\mathbf{W} - \mathbf{R}\|_F^2 : \|\mathbf{W}\|_F = 1\}$$

Proof. Let us denote by P the set of indices corresponding to the p greatest entries (in absolute value) of \mathbf{S} , and by \mathbf{S}_P the matrix we get by setting all the other entries of \mathbf{S} to 0. Let us also introduce $\mathbf{S}_{\bar{P}} = \mathbf{S} - \mathbf{S}_P$, the matrix we get by keeping only the $n^2 - p$ littlest entries. We have:

$$\|\mathbf{S}\|_F^2 = \|\mathbf{S}_P\|_F^2 + \|\mathbf{S}_{\bar{P}}\|_F^2 \quad (3)$$

and

$$\mathbf{S}_P = \arg \min_{\mathbf{V}} \{\|\mathbf{V} - \mathbf{S}\|_F^2 : \|\mathbf{V}\|_0 \leq p\} = T_p(\mathbf{S}) \quad (4)$$

and

$$P = \arg \max_J \{\|\mathbf{S}_J\|_F : \text{card}(J) \leq p\} \quad (5)$$

The operator B_1 is simply the projection onto the Euclidean sphere of radius 1, that is:

$$B_1(\mathbf{X}) = \frac{1}{\|\mathbf{X}\|_F} \mathbf{X}. \quad (6)$$

Let us now denote by I the support of any $\mathbf{U} \in \mathcal{E}$, we have:

$$\|\mathbf{U} - \mathbf{S}\|_F^2 = \|\mathbf{S}_{\bar{I}}\|_F^2 + \|\mathbf{U} - \mathbf{S}_I\|_F^2 \quad (7)$$

so the problem can be rewritten:

$$(\mathbf{S}^*, I^*) = \arg \min_{(\mathbf{U}, I)} \{\|\mathbf{S}_{\bar{I}}\|_F^2 + \|\mathbf{U} - \mathbf{S}_I\|_F^2 : \|\mathbf{U}\|_F = 1, \text{card}(I) \leq p\} \quad (8)$$

Isolating \mathbf{U} we have:

$$\mathbf{S}^* = \arg \min_{\mathbf{U}} \{\|\mathbf{U} - \mathbf{S}_{I^*}\|_F^2 : \|\mathbf{U}\|_F \leq q\} = B_1(\mathbf{S}_{I^*}) \quad (9)$$

with:

$$\begin{aligned} I^* &= \arg \min_I \{\|\mathbf{S}_{\bar{I}}\|_F^2 + \inf_{\mathbf{U}} \{\|\mathbf{U} - \mathbf{S}_I\|_F^2 : \|\mathbf{U}\|_F = 1\} : \text{card}(I) \leq p\} \\ I^* &= \arg \min_I \{\|\mathbf{S}_{\bar{I}}\|_F^2 + \|B_1(\mathbf{S}_I) - \mathbf{S}_I\|_F^2 : \text{card}(I) \leq p\} \\ I^* &= \arg \min_I \{\|\mathbf{S}_{\bar{I}}\|_F^2 + (1 - \frac{1}{\|\mathbf{S}_I\|_F})^2 \|\mathbf{S}_I\|_F^2 : \text{card}(I) \leq p\} \\ I^* &= \arg \min_I \{\underbrace{\|\mathbf{S}_{\bar{I}}\|_F^2 + \|\mathbf{S}_I\|_F^2}_{\|\mathbf{S}\|_F^2} - 2\|\mathbf{S}_I\|_F : \text{card}(I) \leq p\} \\ I^* &= \arg \max_I \{\|\mathbf{S}_I\|_F : \text{card}(I) \leq p\} \\ I^* &= P \end{aligned} \quad (10)$$

So we have: $\mathbf{S}^* = B_1(\mathbf{S}_P) = B_1(T_p(\mathbf{S}))$. \square

2 Lipschitz moduli proof

Let us look at the Lipschitz moduli of the gradient of the smooth part of the objective:

$$\begin{aligned} & \left\| \nabla_{\mathbf{S}_j^i} H(\mathbf{S}_1^{i+1} \dots \mathbf{S}_1 \dots \mathbf{S}_p^i, \lambda^i) - \nabla_{\mathbf{S}_j^i} H(\mathbf{S}_1^{i+1} \dots \mathbf{S}_2 \dots \mathbf{S}_p^i, \lambda^i) \right\|_F \\ &= \left\| \lambda^i \mathbf{L}^T (\lambda^i \mathbf{L} \mathbf{S}_1 \mathbf{R} - \mathbf{X}) \mathbf{R}^T - \lambda^i \mathbf{L}^T (\lambda^i \mathbf{L} \mathbf{S}_2 \mathbf{R} - \mathbf{X}) \mathbf{R}^T \right\|_F \\ &= (\lambda^i)^2 \left\| \mathbf{L}^T \mathbf{L} (\mathbf{S}_1 - \mathbf{S}_2) \mathbf{R} \mathbf{R}^T \right\|_F \\ &= (\lambda^i)^2 \left\| (\mathbf{R} \mathbf{R}^T) \otimes (\mathbf{L}^T \mathbf{L}) \cdot \text{vec}(\mathbf{S}_1 - \mathbf{S}_2) \right\|_2 \\ &\leq (\lambda^i)^2 \left\| (\mathbf{R} \mathbf{R}^T) \otimes (\mathbf{L}^T \mathbf{L}) \right\|_2 \|\mathbf{S}_1 - \mathbf{S}_2\|_F \\ &= (\lambda^i)^2 \|\mathbf{R}\|_2^2 \cdot \|\mathbf{L}\|_2^2 \|\mathbf{S}_1 - \mathbf{S}_2\|_F. \end{aligned} \quad (11)$$

So we can say that the following quantity is a Lipschitz modulus: $L_j(\mathbf{L}, \mathbf{R}, \lambda^i) = (\lambda^i)^2 \|\mathbf{R}\|_2^2 \cdot \|\mathbf{L}\|_2^2$.