Learning on the symmetric group

Jean-Philippe Vert

Google Zürich,
Septembre 29, 2017
Motivation

Personalized Cancer Therapy

1. Molecular Profiling
2. Prognostic Markers
   - Markers predictive of drug sensitivity/resistance
   - Markers predictive of adverse events

https://pct.mdanderson.org
Data

- $X$ gene expression profile of each patient
- $Y$ survival information of each patient
- $n = 10^2 \sim 10^4$
- $p = 2 \times 10^4$
- Goal: learn to predict $Y$ from $X$
- But... where does $X$ come from?
From raw data to $X$

- **Between-sample variability:** batch effect, drift over time, ...
- **Typical pre-processing:** Quantile normalization per sample
- **Only the relative ordering of features** within each sample is used
- **See also:** pictures (Gonzalez and Woods, 2008), MRI scans (Shinohara et al., 2014), speech (Hilger and Ney, 2006)
The symmetric group $S_p$ is the set of permutations of \{1, \ldots, p\}

How to estimate $Y = f(X)$ where $X \in S_p$?
Represent a permutation $x \in S_p$ by the vector of rank $\Phi(x) \in \mathbb{R}^p$.
- this is a particular quantile normalization
- Diffusion kernel over the Cayley’s graph (Kondor and Barbosa, 2010)
  - but complexity $O(p^{2p})$
- Many other data come as permutations (votes, preferences, ...)

Related work
Outline

1. Supervised quantile normalization
2. The Kendall and Mallows kernels
3. Conclusion
Outline

1. Supervised quantile normalization
2. The Kendall and Mallows kernels
3. Conclusion
Joint work with

Marine Le Morvan

https://arxiv.org/abs/1706.00244
Standard full quantile normalization

Typically followed by a predictive model $f(X)$ on the normalized data
How to choose a "good" target distribution?
Notations

- \( x_1, \ldots, x_n \in \mathbb{R}^p \) a set of \( p \)-dimensional samples

- \( f \in \mathbb{R}^p \) a non-decreasing target distribution (CDF)

- For \( x \in \mathbb{R}^p \), let \( \Phi_f(x) \in \mathbb{R}^p \) be the data after QN with target distribution \( f \)
Standard approaches: learn model after QN preprocessing:

1. Fix $f$ arbitrarily
2. QN all samples to get $\Phi_f(x_1), \ldots, \Phi_f(x_n)$
3. Learn a generalized linear model $(w, b)$ on normalized data:

$$
\min_{w, b} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( w^\top \Phi_f(x_i) + b \right) + \lambda \Omega(w) \right\}
$$

SUQUAN: jointly learn $f$ and $(w, b)$:

$$
\min_{w, b, f} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( w^\top \Phi_f(x_i) + b \right) + \lambda \Omega(w) + \gamma \Omega_2(f) \right\}
$$
SUQAN as matrix regression (1/2)

- For $x \in \mathbb{R}^p$, let $\Pi_x \in \mathbb{R}^{p \times p}$ the permutation matrix of $x$’s entries:

  \[ [\Pi_x]_{ij} = 1 \quad (x_j \text{ is the } i\text{-th smallest feature}) \]

- Quantile normalized $x$ with target distribution $f$ is:

  \[ \Phi_f(x) = \Pi_x f \]

- Example:

  \[
  x = \begin{pmatrix}
    4.5 \\
    1.2 \\
    10.1 \\
    8.9
  \end{pmatrix}
  \quad
  \Pi_x = \begin{pmatrix}
    0 & 1 & 0 & 0 \\
    1 & 0 & 0 & 0 \\
    0 & 0 & 0 & 1 \\
    0 & 0 & 1 & 0
  \end{pmatrix}
  \quad
  f = \begin{pmatrix}
    0 \\
    1 \\
    3 \\
    4
  \end{pmatrix}
  \]

  \[
  \Phi_f(x) = \Pi_x f = \begin{pmatrix}
    1 \\
    0 \\
    4 \\
    3
  \end{pmatrix}
  \]
SUQAN as matrix regression (2/2)

- SUQUAN solves

\[
\min_{w, b, f} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( w^\top \Phi_f(x_i) + b \right) + \lambda \Omega(w) + \gamma \Omega_2(f) \right\}
\]

\[
= \min_{w, b, f} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell \left( w^\top \Pi_{x_i} f + b \right) + \lambda \Omega(w) + \gamma \Omega_2(f) \right\}
\]

\[
= \min_{w, b, f} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell \left( < w f^\top, \Pi_{x_i} >_F + b \right) + \lambda \Omega(w) + \gamma \Omega_2(f) \right\}
\]

- A particular rank-1 matrix optimization, \( x \) is replaced by \( \Pi_x \)
- Non-convex
- Local optimum found by alternatively optimizing \( f \) and \( w \)
Constraints on $f$

- **Ridge**

\[
\mathcal{F}_0 = \left\{ f \in \mathbb{R}^p : \frac{1}{p} \sum_{i=1}^{p} f_i^2 \leq 1 \right\}.
\]

- **Non-decreasing**

\[
\mathcal{F}_{\text{BND}} = \mathcal{F}_0 \cap \mathcal{I}_0, \quad \text{where} \quad \mathcal{I}_0 = \left\{ f \in \mathbb{R}^p : f_1 \leq f_2 \leq \ldots \leq f_p \right\}
\]

- **Non-decreasing and smooth**

\[
\mathcal{F}_{\text{SPAV}} = \left\{ f \in \mathcal{I}_0 : \sum_{j=1}^{p-1} (f_{j+1} - f_j)^2 \leq 1 \right\}.
\]
SUQUAN-BND and SUQUAN-PAVA

5.2 SUQUAN-BND and SUQUAN-SPAV

We now focus on approximate algorithms to solve (8) in the case where \( F = F_{\text{BND}} \) or \( F = F_{\text{SPAV}} \). We then compare four methods to estimate \( w \) from \( n \) observations:

6

**Algorithm 2: SUQUAN-BND and SUQUAN-SPAV**

**Input:** \((x_1, y_1), \ldots, (x_n, y_n), f_{\text{init}} \in \mathcal{I}_0, \lambda \in \mathbb{R}\)

**Output:** \( f \in \mathcal{I}_0 \) target quantile

1. **for** \( i = 1 \) to \( n \) **do**
   2. \( \text{rank}_i, \text{order}_i \leftarrow \text{sort}(x_i) \)
   3. **end for**

4. \( w, b \leftarrow \arg\min_{w,b} \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( w^\top f_{\text{init}}[\text{rank}_i] + b \right) + \lambda ||w||^2 \)
   (standard linear model optimisation)

5. \( f \leftarrow \arg\min_{f \in \mathcal{F}_{\text{BND}}} \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( f^\top w[\text{order}_i] + b \right) \)
   (isotonic optimisation problem using PAVA as prox)
   OR
   \( f \leftarrow \arg\min_{f \in \mathcal{F}_{\text{SPAV}}} \frac{1}{n} \sum_{i=1}^{n} \ell_i \left( f^\top w[\text{order}_i] + b \right) \)
   (smoothed isotonic optimisation problem using SPAV as prox)

- Alternate optimization in \( w \) and \( f \), monotonicity constraint on \( f \)
- Accelerated proximal gradient optimization for \( f \), using the Pool Adjacent Violators Algorithm (PAVA, Barlow et al. (1972)) or the Smoothed Pool Adjacent Violators algorithm (SPAV, Sysoev and Burdakov (2016)) as proximal operator.
A variant: SUQUAN-SVD

<table>
<thead>
<tr>
<th>Algorithm 1: SUQUAN-SVD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
</tr>
<tr>
<td>$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^p \times {-1, 1}$</td>
</tr>
<tr>
<td><strong>Output:</strong> $f \in F_0$ target quantile</td>
</tr>
<tr>
<td>1: $M_{LDA} \leftarrow 0 \in \mathbb{R}^{p \times p}$</td>
</tr>
<tr>
<td>2: $n_+ \leftarrow</td>
</tr>
<tr>
<td>3: $n_- \leftarrow</td>
</tr>
<tr>
<td>4: for $i = 1$ to $n$ do</td>
</tr>
<tr>
<td>5: Compute $\Pi_{x_i}$ (by sorting $x_i$)</td>
</tr>
<tr>
<td>6: $M_{LDA} \leftarrow M_{LDA} + \frac{y_i}{n_y} \Pi_{x_i}$</td>
</tr>
<tr>
<td>7: end for</td>
</tr>
<tr>
<td>8: $(\sigma, w, f) \leftarrow \text{SVD}(M_{LDA}, 1)$</td>
</tr>
</tbody>
</table>

- Ridge penalty (no monotonicity constraint), equivalent to rank-1 regression problem
- SVD finds the closest rank-1 matrix to the LDA solution:

$$M_{LDA} = \frac{1}{n_+} \sum_{i : y_i = +1} \Pi_{x_i} - \frac{1}{n_-} \sum_{i : y_i = +1} \Pi_{x_i}$$

- Complexity $O(np \ln(p))$ (same as QN only)
## Experiments: Simulations

- True distribution of $X$ entries is normal
- Corrupt data with a cauchy, exponential, uniform or bimodal gaussian distributions.
- $p = 1000$, $n$ varies, logistic regression.

### Table

<table>
<thead>
<tr>
<th>Number of samples</th>
<th>AUC</th>
<th>Log. reg. original</th>
<th>Log. reg. corrupted</th>
<th>SUQUAN BND</th>
<th>SUQUAN SPAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Figure

- **AUC:** Comparison of different models across varying numbers of samples.
- **Euclidean distance to original target quantile:** Shows the change in distance as the number of samples increases.
Experiments: CIFAR-10

- Example: horse vs. plane
- Different methods learn different quantile distributions

![Images of horse vs. plane with different methods applied]

<table>
<thead>
<tr>
<th>Original</th>
<th>Median</th>
<th>SVD</th>
<th>SUQUAN BND</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="original" alt="Images" /></td>
<td><img src="median" alt="Images" /></td>
<td><img src="SVD" alt="Images" /></td>
<td>![Images](SUQUAN BND)</td>
</tr>
</tbody>
</table>

![Graphs of quantile distributions]

0 400 800 0 400 800 0 400 800
Experiments: CIFAR-10

- Image classification into 10 classes (45 binary problems)
- \( n = 5,000 \) per class, \( p = 1,024 \) pixels
Experiments: gene expression data

- Breast cancer prognosis from gene expression data.
  - \( X = \) expression levels of 22,283 genes of the tumour at diagnosis
  - \( Y = 1 \) if cancer relapse within 6 years of diagnosis, 0 otherwise

- 4 datasets:

<table>
<thead>
<tr>
<th>DATASET NAME</th>
<th># PATIENTS</th>
<th># POSITIVES</th>
<th>% POSITIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE1456</td>
<td>141</td>
<td>37</td>
<td>0.26</td>
</tr>
<tr>
<td>GSE2034</td>
<td>271</td>
<td>104</td>
<td>0.38</td>
</tr>
<tr>
<td>GSE2990</td>
<td>106</td>
<td>32</td>
<td>0.30</td>
</tr>
<tr>
<td>GSE4922</td>
<td>225</td>
<td>73</td>
<td>0.32</td>
</tr>
</tbody>
</table>
# Results: gene expression data

<table>
<thead>
<tr>
<th></th>
<th>RAW</th>
<th>RMA</th>
<th>CAUCHY</th>
<th>EXP.</th>
<th>UNIF.</th>
<th>GAUS.</th>
<th>MEDIAN</th>
<th>SVD</th>
<th>BND</th>
<th>SPAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE1456</td>
<td>65.94</td>
<td>68.73</td>
<td>59.56</td>
<td>68.86</td>
<td>68.72</td>
<td>69.00</td>
<td>69.06</td>
<td>57.60</td>
<td>71.44</td>
<td>69.60</td>
</tr>
<tr>
<td>GSE2034</td>
<td>74.52</td>
<td>75.42</td>
<td>61.91</td>
<td>74.53</td>
<td>75.22</td>
<td>76.45</td>
<td>74.92</td>
<td>52.61</td>
<td>70.50</td>
<td>76.11</td>
</tr>
<tr>
<td>GSE2990</td>
<td>57.01</td>
<td>60.43</td>
<td>54.72</td>
<td>61.25</td>
<td>56.25</td>
<td>58.66</td>
<td>59.72</td>
<td>52.51</td>
<td>59.22</td>
<td>59.94</td>
</tr>
<tr>
<td>GSE4922</td>
<td>58.52</td>
<td>58.86</td>
<td>55.24</td>
<td>58.81</td>
<td>55.66</td>
<td>60.01</td>
<td>59.18</td>
<td>52.39</td>
<td>61.82</td>
<td>61.41</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>64.00</td>
<td>65.86</td>
<td>57.86</td>
<td>65.86</td>
<td>63.96</td>
<td>66.03</td>
<td>65.72</td>
<td>53.78</td>
<td>65.75</td>
<td>66.77</td>
</tr>
</tbody>
</table>

![GSE2034](image1.png)  
![GSE4922](image2.png)
Estimated distribution: iteration=0
Estimated distribution: iteration=1
Estimated distribution: iteration=2
1. Supervised quantile normalization

2. The Kendall and Mallows kernels

3. Conclusion
Joint work with

Yunlong Jiao

https://hal.archives-ouvertes.fr/hal-01279273
An idea: all pairwise comparisons

Replace $x \in \mathbb{R}^p$ by $\Phi(x) \in \{0, 1\}^{p(p-1)/2}$:

$$
\Phi_{i,j}(x) = \begin{cases} 
1 & \text{if } x_i \leq x_j, \\
0 & \text{otherwise.}
\end{cases}
$$
Related work: Top scoring pairs (TSP)

(Geman et al., 2004; Tan et al., 2005; Leek, 2009)
Practical challenge

- Need to store $O(p^2)$ bits per sample
- Need to train a model in $O(p^2)$ dimensions
### Theorem (Wahba, Schölkopf, …)

Training a linear model over a representation $\Phi(x) \in \mathbb{R}^Q$ of the form:

$$
\min_{w \in \mathbb{R}^Q} \frac{1}{n} \sum_{i=1}^{n} \ell(w^\top \Phi(x_i), y_i) + \lambda \|w\|^2
$$

can be done efficiently, independently of $Q$, if the kernel

$$
K(x, x') = \Phi(x)^\top \Phi(x')
$$

can be computed efficiently.

Ex: ridge regression, $O(Q^3 + nQ^2)$ becomes $O(n^3 + n^2 T)$

Other: SVM, logistic regression, Cox model, survival SVM, …
Kernel trick for us: Kendall’s $\tau$

$$\Phi(x)^\top \Phi(x') = \tau(x, x') \text{ (up to a scaling)}$$

Good news for SVM and kernel methods!
More formally

- For two permutations $\sigma, \sigma'$ let $n_c(\sigma, \sigma')$ (resp. $n_d(\sigma, \sigma')$) the number of concordant (resp. discordant) pairs.
- The Kendall kernel (a.k.a. Kendall tau coefficient) is defined as
  \[ K_\tau(\sigma, \sigma') = \frac{n_c(\sigma, \sigma') - n_d(\sigma, \sigma')}{\binom{p}{2}}. \]
- The Mallows kernel is defined for any $\lambda \geq 0$ by
  \[ K_\lambda^M(\sigma, \sigma') = e^{-\lambda n_d(\sigma, \sigma')} . \]

**Theorem (Jiao and V., 2015)**
The Kendall and Mallows kernels are positive definite.

**Theorem (Knight, 1966)**
These two kernels for permutations can be evaluated in $O(p \log p)$ time.
Kondor and Barbarosa (2010) proposed the diffusion kernel on the Cayley graph of the symmetric group generated by adjacent transpositions.

Computationally intensive \((O(p^{2p}))\)

Mallows kernel is written as

\[
K_M^λ(\sigma, \sigma') = e^{-\lambda n_d(\sigma, \sigma')},
\]

where \(n_d(\sigma, \sigma')\) is the shortest path distance on the Cayley graph.

It can be computed in \(O(p \log p)\)
Application: supervised classification

Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>No. of features</th>
<th>No. of samples (training/test)</th>
<th>$C_1$</th>
<th>$C_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer 1</td>
<td>23624</td>
<td>44/7 (Non-relapse)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer 2</td>
<td>22283</td>
<td>142 (Non-relapse)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer 3</td>
<td>22283</td>
<td>71 (Poor Prognosis)</td>
<td>138 (Good Prognosis)</td>
<td></td>
</tr>
<tr>
<td>Colon Tumor</td>
<td>2000</td>
<td>40 (Tumor)</td>
<td></td>
<td>22 (Normal)</td>
</tr>
<tr>
<td>Lung Cancer 1</td>
<td>7129</td>
<td>24 (Poor Prognosis)</td>
<td>62 (Good Prognosis)</td>
<td></td>
</tr>
<tr>
<td>Lung Cancer 2</td>
<td>12533</td>
<td>16/134 (ADCA)</td>
<td>16/15 (MPM)</td>
<td></td>
</tr>
<tr>
<td>Medulloblastoma</td>
<td>7129</td>
<td>39 (Failure)</td>
<td></td>
<td>21 (Survivor)</td>
</tr>
<tr>
<td>Ovarian Cancer</td>
<td>15154</td>
<td>162 (Cancer)</td>
<td></td>
<td>91 (Normal)</td>
</tr>
<tr>
<td>Prostate Cancer 1</td>
<td>12600</td>
<td>50/9 (Normal)</td>
<td></td>
<td>52/25 (Tumor)</td>
</tr>
<tr>
<td>Prostate Cancer 2</td>
<td>12600</td>
<td>13 (Non-relapse)</td>
<td></td>
<td>8 (Relapse)</td>
</tr>
</tbody>
</table>

Methods

- Kernel machines Support Vector Machines (SVM) and Kernel Fisher Discriminant (KFD) with Kendall kernel, linear kernel, Gaussian RBF kernel, polynomial kernel.
- Top Scoring Pairs (TSP) classifiers Tan et al. (2005).
- Hybrid scheme of SVM + TSP feature selection algorithm.
Results

Kendall kernel SVM

- **Competitive accuracy!**
- Less sensitive to regularization parameter!
- No need for feature selection!
Results

Kendall kernel SVM
- Competitive accuracy!
- Less sensitive to regularization parameter!
- No need for feature selection!
Kendall kernel SVM
- Competitive accuracy!
- Less sensitive to regularization parameter!
- No need for feature selection!
Application: clustering

- APA data (full rankings)
- \( n = 5738, p = 5 \)
- (new) Kernel k-means vs (standard) k-means in \( S_5 \)
- Show silhouette as a function of number of clusters (higher better)
Extension to partial rankings

- Two interesting types of partial rankings are interleaving partial ranking
  \[ x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_k}, \quad k \leq n. \]
  and top-k partial ranking
  \[ x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_k} \succ X_{\text{rest}}, \quad k \leq n. \]

- Partial rankings can be uniquely represented by a set of permutations compatible with all the observed partial orders.

**Theorem**

For these two particular types of partial rankings, the convolution kernel (Haussler, 1999) induced by Kendall kernel

\[
K^*_T(R, R') = \frac{1}{|R||R'|} \sum_{\sigma \in R} \sum_{\sigma' \in R'} K_T(\sigma, \sigma')
\]

can be evaluated in \(O(k \log k)\) time.
Two interesting types of partial rankings are **interleaving partial ranking**

\[ x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_k}, \quad k \leq n. \]

and **top-\(k\) partial ranking**

\[ x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_k} \succ X_{\text{rest}}, \quad k \leq n. \]

Partial rankings can be **uniquely represented** by a set of permutations compatible with all the observed partial orders.

**Theorem**

*For these two particular types of partial rankings, the convolution kernel (Haussler, 1999) induced by Kendall kernel*

\[
K^*_\tau(R, R') = \frac{1}{|R||R'|} \sum_{\sigma \in R} \sum_{\sigma' \in R'} K_\tau(\sigma, \sigma')
\]

*can be evaluated in \(O(k \log k)\) time.*
Instead of $\Phi : \mathbb{R}^p \rightarrow \{0, 1\}^{p(p-1)/2}$, consider the continuous mapping $\Psi_a : \mathbb{R}^p \rightarrow \mathbb{R}^{p(p-1)/2}$:

$$
\Psi_a(x) = E\Phi(x + \epsilon) \quad \text{with} \quad \epsilon \sim (U[-\frac{a}{2}, \frac{a}{2}])^n
$$

Corresponding kernel $G_a(x, x') = \Psi_a(x)\top\Psi_a(x')$
Computation of $G(x, x')$

- $G_a(x, x')$ can be computed exactly in $O(p^2)$ by explicit computation of $\Psi_a(x)$ in $\mathbb{R}^{p(p-1)/2}$

- $G_a(x, x')$ can be computed approximately in $O(D^2p \log p)$ by Monte-Carlo approximation:

$$\tilde{G}_a(x, x') = \frac{1}{D^2} \sum_{i,j=1}^{D} K(x + \epsilon_i, x' + \epsilon'_j)$$

- Theorem: for supervised learning, Monte-Carlo approximation is better\(^1\) than exact computation when $n = o(p^{1/3})$

\(^1\)faster for the same accuracy
Performance of $G_a(x, x)$

![Graph showing the performance of $G_a(x, x)$ for different noise window sizes. The x-axis represents the noise window size $a$, and the y-axis represents $cvacc$. The graph includes lines for different approximations of $G_a(x, x)$, such as SVMkdtALLalt−−MCapprox (D=1), SVMkdtALLalt−−MCapprox (D=3), SVMkdtALLalt−−MCapprox (D=5), SVMkdtALLalt−−MCapprox (D=7), and SVMkdtALLalt−−MCapprox (D=9).]
Outline

1. Supervised quantile normalization
2. The Kendall and Mallows kernels
3. Conclusion
**Conclusion**

- Representing omics data as **permutations** has some potential
  - **Kendall and Mallows** kernel in \( O(p \ln(p)) \)
  - **SUQUAN** supervised quantile normalization as matrix regression

**Ongoing work:**
- Extension of **SUQUAN** to nonlinear models (neural nets..)
- Extention of **SUQUAN** to Kendall representation (weighted Kendall correlation...)
1/ PROMOTION DU PROGRAMME

- Rencontrez avec les directeurs, directeurs scientifiques et directeurs des études de l'ENS, Mines, ENSCP, ESPCI pour organiser la promotion du programme auprès des étudiants.

- Ecole des Mines : présentation d'ITI devant le conseil de l'enseignement ; 8 représentants d'élèves

- ENSCP : mailing de présentation d'ITI aux 3A via la direction des études.

- ENS : mailing ; présentation directe auprès des étudiants (2 élèves présents) ; diffusion des plaquettes et du syllabus

- ESPCI : mailing aux 3A ; présentation d'ITI / rencontre avec les étudiants en présence de la directrice de la scolarité et directrice des relations ; 8 élèves présents

- Contacts en cours avec les Ponts et l'ENS Lyon (en attente de réponse)

- Rencontrez à l'ANRT avec le délégué général et la chef du service CIFRE : accord de communication sur le programme ITI via le site intern et de l'ANRT (rubrique "zoom sur")

- Promotion aux Réunions Universités Entreprises (RUE)

- Echange avec Stéphane Mallat et réflexion sur la pertinence du programme tronc commun

- Rencontre avec le responsable des relations internationales de la National Taiwan University à l'ESPCI : présentation d'ITI

- Contact en cours pour visite d'entreprise

- Article dans les Echos

- Création du logo PSL / ITI et du Diplôme Supérieur de Recherche et d'Innovation de PSL / ITI (DSRI), dépôt INPI en cours.

2/ MISE EN ŒUVRE OPERATIONNELLE DU PROGRAMME
References


