Statistical Learning

Prof. Dr. Gilles Blanchard, Dr. Alexandra Carpentier, Dr. Jana de Wiljes, Prof. Dr. Markus Reiss

Introduction (German)

Maschinelles/statistisches Lernen ist ein schnell wachsender wissenschaftlicher Bereich mit vielen Verknüpfungen zur Mathematik, insbesondere zur mathematischen Statistik. Ziel des Seminars ist das Verständnis statistischer Modelle und Begriffe, die grundlegend sind fr die mathematische Analyse des maschinelles Lernens. Im ersten Teil behandeln wir den Zusammenhang zwischen empirischer Risikominimierung und empirischen Prozessen sowie den fundamentalen Begriff der Regularisierung komplexer Modelle. Im zweiten Teil werden Methoden des sogenannten Aktiv- und Online-Lernens studiert. Das Seminar strebt nicht eine vollstndige Prsentation von Themen und Ergebnissen des maschinelles Lernens an, sondern fokussiert auf besonders interessante Aspekte. Leitlinie ist die mathematisch fundierte Beschreibung und Analyse wichtiger Konzepte des modernen statistischen Lernens.

List of projects

1. Project 1 : Concentration (and complexity) Tommaso Rosati
2. Project 2 : Complexity (and concentration) Matthias Kirchler
3. Project 3 :Classification with support vector machines Benjamin Hartmann oder Christopher Lennan
4. Project 4 : The Lasso, a specific example of regularisation Benjamin Hartmann oder Christopher Lennan
5. Project 5 : Kernel regression and RKHS Mathäus Deutsch
6. Project 6 : Random projections and the Johnson Lindenstrauss Lemma Jonas Lieber und Florian Hildebrandt
7. Project 7 : Prediction with expert advices Deindra Haag
8. Project 8 : The stochastic bandit problem - upper bounds Lorenz Richter oder Katharina Merz
9. Project 9 : The stochastic bandit problem - problem independent lower bound Natalia Walko
10. Project 10 : The adversarial bandit problem Lorenz Richter oder Katharina Merz
11. Project 11 : Infinitely many armed bandits without structure Matti Weigel
12. Project 12 : Infinitely many armed bandits with structure - Optimisation using bandits Philipp Trunschke

Contents

1. Batch Learning

1.1. Foundations of Empirical Risk Minimization (ERM) : concentration and complexity

A very general and common theoretical objective in statistical learning is to study the behaviour of empirical risk minimizers.

Let $\mathcal{D}_n = (D_1, \dots, D_n)$ be a dataset of n points, generated in an i.i.d, fashion according to a distribution μ^*. We will write μ_n for the empirical distribution induced by $\mathcal{D}_n$, i.e. for any μ^*-measurable set A, we set

$$\mu_n(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{D_i \in A\}.$$

Let $\mathcal{P}$ be a class of models. These models can be anything : distributions, classifiers, tests....

Let $L(\mu, \theta)$ be a risk function, i.e. a function from the set of distributions (whatever this means...) and $\mathcal{P}$ to $\mathbb{R}$. We will be interested in the *risk minimiser*

$$L^* = \inf_{\theta \in \mathcal{P}} L(\mu^*, \theta).$$

i.e. *minimum of the loss function.* In general the loss function is taken such that it has values in $\mathbb{R}^+$, but not necessarily.

Examples : Classification and regression

Now in general one does not know μ^* so this minimum is unreachable....

(Plug-in) Idea : use μ_n instead of μ^*.

So we will be interested in the *empirical risk minimiser*

$$L_n = \min_{\theta \in \mathcal{P}} L(\mu_n, \theta).$$

We then want to control the *generalisation gap*

$$R_n = \sup_{\theta} |L(\mu_n, \theta) - L(\mu^*, \theta)|.$$

Note that in particular

$$|L_n - L^*| \le R_n.$$

Remark also that if we have a bound on that, we also have a bound on the *generalisation error* of any model element selected by ERM, i.e. we have a bound on how the chosen ERM would perform for a new data point.

In order to do that, one should first remark that

$$\begin{aligned} R_n &\leq \sup_{\theta \in \mathcal{P}} |L(\mu_n, \theta) - L(\mu^*, \theta)| \\ &\leq \sup_{\theta \in \mathcal{P}} |L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta)| + \sup_{\theta \in \mathcal{P}} |\mathbf{E}L(\mu_n, \theta) - L(\mu^*, \theta)| \\ &= \sup_{\theta \in \mathcal{P}} |L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta)| + B(\mathcal{P}, L, \mu^*), \end{aligned}$$

where $B(\mathcal{P}, \mu^*)$ is fixed and depends only on $\mathcal{P}$, L and μ^* but not on the data (it is often 0 for instance for summable losses). What remains to do is then to bound with high probability the *empirical process*

$$\sup_{\theta \in \mathcal{P}} |L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta)| := \sup_{\theta \in \mathcal{P}} |Z_\theta|.$$

Hope : When n is large, $L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta) = Z_\theta$ goes to 0 - so when n is large, $\sup_{\theta \in \mathcal{P}} |L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta)| = \sup_{\theta \in \mathcal{P}} |Z_\theta|$ goes to 0.

Indeed, if n is very large, then (under the LLN conditions)

$$Z_\theta = L(\mu_n, \theta) - \mathbf{E}L(\mu_n, \theta) \rightarrow_{a.s.} 0.$$

So obviously if there is a finite number of elements in $\mathcal{P}$

$$\sup_{\theta \in \mathcal{P}} |Z_\theta| \rightarrow_{a.s.} 0.$$

But is this always enough to consider asymptotics? And what if $\mathcal{P}$ is not finite?

Example in the regression setting : Consider a function θ^* that is defined over a $1/n$ grid of $[0, 1]$. Assume that we observe noisy evaluation of f as $Y_i = \theta^*(X_i) + \epsilon_i$ where the ϵ_i are i.i.d. and a white noise and the X_i are i.i.d. and uniformly distributed over the grid. Here the data $D_i = (Y_i, X_i)$ are i.i.d. and distributed according to μ^* which depends on θ^* and the distribution of the noise.

We want to estimate θ^* and in this respect, we write the risk as

$$L(\mu, \theta) = \mathbb{E}_{(X,Y) \sim \mu}(Y - \theta(X))^2.$$

We have

$$L(\mu_n, \theta) = \frac{1}{n} \sum_i (Y_i - \theta(X_i))^2.$$

So we have

$$\begin{aligned}\mathbb{E}_{\mu^*}\sup_{\theta\in\mathcal{P}}|Z_\theta| &\geq \mathbb{E}_{\mu^*}\sup_{\theta\in\mathcal{P}} Z_\theta\\ &= \mathbb{E}_{\mu^*}\sup_{\theta\in\mathcal{P}}\Bigg(L(\mu_n,\theta)-\mathbf{E}_{\mu^*}L(\mu_n,\theta)\Bigg)\\ &= \mathbb{E}_{\mu^*}\sup_{\theta\in\mathcal{P}}\Bigg(\frac{1}{n}\sum_i\Big(\big(\theta^*(X_i)-\theta(X_i)+\epsilon_i\big)^2-\mathbb{E}_{\mu^*}\big(\theta^*(X_i)-\theta(X_i)+\epsilon_i\big)^2\Big)\Bigg)\\ &= \mathbb{E}_{\mu^*}\sup_{\theta\in\mathcal{P}}\frac{2}{n}\sum_{i=1}^n\epsilon_i(\theta^*(X_i)-\theta(X_i)).\end{aligned}$$

Consider the special case $\theta^*=0$ and $\epsilon_i\sim\mathcal{R}$. What is the order of magnitude in the case where $\mathcal{P}$ contains just one element of $\{0,1\}^n$? What happens if $\mathcal{P}$ contains all elements of $\{-1,1\}^n$?

So in other words, if $\mathcal{P}$ is complex, and if n is not "large enough" with respect to the complexity of $\mathcal{P}$, problems can occur...

Question : Can we take the complexity of $\mathcal{P}$ into account in the convergence results?

In order to do that, we will try to develop convergence results that are non-asymptotic and hold for any n. This theory is called *concentration*. We can then try to apply these non-asymptotic results to all elements of $\mathcal{P}$ *simultaneously.*

Global objective : We want to bound with high probability the empirical risk. In order to do that, we want to bound with high probability

$$\sup_{\theta\in\mathcal{P}}|L(\mu_n,\theta)-\mathbf{E}L(\mu_n,\theta)|=\sup_{\theta\in\mathcal{P}}|Z_\theta|,$$

depending on n and on the complexity of $\mathcal{P}$.

Project 1 : Concentration (and complexity)

Objectives : In order to bound

$$\sup_{\theta\in\mathcal{P}}|Z_\theta|,$$

a first objective is to derive a high probability bound, for a finite n, and first a fixed θ, on

$$|Z_\theta|.$$

The assumptions that are made on the loss function L so that concentration happens should be discussed. Then the case of a finite set $\mathcal{P}$ will be discussed, and a bound on

$$\sup_{\theta \in \mathcal{P}} |Z_\theta|,$$

will then be presented. This project shall focus in priority to the concentration of measure phenomenon and slightly less on the measures of complexity - the set $\mathcal{P}$ is assumed to be finite although potentially large and not necessarily negligible with respect to n. The tasks are the following.

1. Present and provide a proof of *Hoeffding's inequality* (this is done in e.g. Chapter 1 in the book Gyorfi (2002)). Explain how this can help for bounding Z_θ in the case where the specific shape of the loss function is
$$L(\mu_n, \theta) = \frac{1}{n}\sum_{i=1}^n \theta(X_i),$$
where θ is a function defined on the domain of μ_* and is bounded by 1.
2. Present the proof of a general concentration inequality (like the *bounded difference concentration inequality* or *Mc Diarmid concentration inequality*) that can be applied to a more general ERM function. This can be found for instance in Chapter 1.3 in the book Gyorfi (2002). It is also interesting to look at the book Boucheron et.al (2013) for a more global perspective.
3. Consider the case where the cardinal of $\mathcal{P}$ is bounded by p, and present a bound (*union bound*) on
$$\sup_{\theta \in \mathcal{P}} |Z_\theta|,$$
in order to highlight how p impacts the concentration bound.

General material at ERM can be found both in Chapter 1 in the book Gyorfi (2002), and also in the two first chapters of the book Mohri et.al (2012).
References : Chapter 1 in the book Gyorfi (2002) (contains similar material but formulated a bit differently). Optionally, one can take a look at the book Boucheron et.al (2013).

Project 2 : Complexity (and concentration)

Objective : We can decompose the target in a bias plus a variation term as

$$\begin{aligned}
&\sup_\theta |Z_\theta| \\
&\leq \mathbb{E}\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)| + \Big|\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)| - \mathbb{E}\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)|\Big| \\
&= B + D.
\end{aligned}$$

One then wants to bound these terms separately. Here the accent will be put more on understanding how to bound the non-stochastic quantity

$$B = \mathbb{E}\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)|,$$

and how the complexity of $\mathcal{P}$ will express itself. But bounds on the variation term

$$V = \Big|\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)| - \mathbb{E}\sup_\theta |L(\mu_n,\theta) - \mathbb{E}L(\mu_n,\theta)|\Big|$$

will also be presented - no proofs will be provided.

The whole project will be done in the *binary classification setting*. The data $\mathcal{D}_n$ are of the form $\mathcal{D}_n = (D_i)_i = (X_i, Y_i)_i$ where X_i are the points and the Y_i are the binary labels. $\mathcal{P}$ is a set of binary classifiers, and the loss function L is

$$L(\mu,\theta) = \mathbb{E}_{(X,Y)\sim\mu}(\mathbf{1}_{\theta(X)=Y}).$$

One has for the empirical loss

$$L(\mu_n,\theta) = \frac{1}{n}\sum_i \mathbf{1}_{\theta(X_i)=Y_i}).$$

See e.g. Chapter 1 in the book Gyorfi (2002).

The tasks are the following.

1. Use the *expected Rademacher complexity* to bound B. This can be found in the book Gyorfi (2002), Chapter 1.4, and also in the book Mohri et.al (2012), chapters 3.1 and 3.2 and 3.3.
2. One then has to *bound the expected Rademacher complexity.*
   - First it will be assumed that $\mathcal{P}$ has a finite *Vapnik Chervonenkis dimension.* This can be found in the book Gyorfi (2002), Chapter 1.4, and also in the book Mohri et.al (2012), chapters 3.1 and 3.2 and 3.3.
   - Then in a more general case one might want to apply *chaining* to bound expected Rademacher complexity (see the book Massart (2007), P184).
3. Understand how concentration inequalities like *Talagrand's inequality* can be used to bound the deviation term V - no proofs needed.

References : The book Mohri et.al (2012), chapters 3.1, 3.2 and 3.3. Also Chapter 1 in the book Gyorfi (2002) (contains similar material but formulated a bit differently) and the book Massart (2007), P184.

1.2. Specific examples of learning problems and the notion of penalisation

As we have seen, the complexity of the model $\mathcal{P}$ impacts empirical risk minimisation, and the generalisation error through the generalisation gap. For this reason, one would like to take $\mathcal{P}$ as small as possible but also such that L^* is small - i.e. the risk minimiser would still have a good performance on it.

Quite often, one cannot assume that the "smallest possible but good" model $\mathcal{P}$ is known before hand - one does not have enough prior knowledge and therefore one does not want to make too many model assumptions. As seen before, the complexity of the model will appear in ERM, and be a problem for estimation.

In many cases however, it is not absurd to assume that there is a nested class of model $\mathcal{P}(k)$ - where $\mathcal{P}(k) \subset \mathcal{P}(k')$ for $k \leq k'$, and the smallest model is very small, while the largest model is huge. For each model, one can do ERM and obtain an estimator. Now the question is the following.

Question : How can we choose among all estimators computed on the nested models $\mathcal{P}(k)$?

This question aims in some sense at trying to learn the model k which is the smallest, and where the data still "fits" well. Imagine that we have a high probability bound $V_n(k)$ on the generalisation gap on each model $\mathcal{P}(k)$, then the generalisation error of the ERM is bounded as

$$L^*(k) + V_n(k),$$

where $L^*(k)$ is the risk minimizer on $\mathcal{P}(k)$. So a "smallest possible but good" model therefore minimises $L^*(k)+V_n(k)$: one wants to *penalise* for the complexity of the model k - keeping in mind that the larger k, the worst the deviations $V(k)$, but the smaller the "distance" to the model $L^*(k)$.

Since $L^*(k)$ is unknown, a good idea is to penalise the ERM in k namely $L_n(k)$ as

$$L_n(k) + pen(k),$$

where $pen(k)$ is a penalty for using model k - and this penalty is an *increasing* function of the complexity k - typically we would like it to be of order $V_n(k)$. This penalty term penalises the use of a more complex model and prevents its use if the gains of using it are not compensating the penalty, roughly speaking.

1.2.1. Classification

A specific example of ERM is the binary classification setting. Here the data $\mathcal{D}_n$ are of the form $(X_i, Y_i)_i$ where the X_i are the points sampled in an i.i.d. fashion on the domain, and the

$$Y_i \sim \mathcal{B}(\theta^*(X_i)),$$

are independent Bernoulli random variables and where θ^* is a function between 0 and 1 defined on the domain of the X. $\mathcal{P}$ is a set of binary classifiers, and the loss function L is

$$L(\mu, \theta) = \mathbb{E}_{(X,Y)\sim\mu}(\mathbf{1}_{\theta(X)=Y}).$$

One has for the empirical loss

$$L(\mu_n, \theta) = \frac{1}{n}\sum_i \mathbf{1}_{\theta(X_i)=Y_i}).$$

This setting is an ERM setting as posed before.The main question is on : indeed, one minimises the empirical risk often in order to determine a suitable model element θ with respect to the data. In many cases, there are many such θ that minimise the risk.

Question : How can we decide between many parameters that minimise, or approximately minimise, the empirical risk which one is the most appropriate?

This will be linked to the penalisation idea introduced earlier.

Project 3 :Classification with Support Vector Machines (SVM)

Objective : A natural set of classifiers that one can want to consider are *linear classifiers*, i.e.

$$\mathcal{P} = \{\theta = (w, b) \in \mathbb{R}^n \times \mathbb{R}, \theta(x) = \mathrm{sign}(\langle w, x\rangle + b)\}.$$

Now in many cases, there are many elements of $\mathcal{P}$ that minimize the empirical risk. Then one wants to have a good rule for finding the most adequate separator. A good rule for doing so in the *separable case* (i.e. when the two classes can be perfectly separated) is to choose the classifier that corresponds to an hyperplan that maximizes the *margin*, i.e. the distance of the two classes to the hyperplan - this defines an interesting optimisation problem. In the *non-separable* case, one has to relax the margin assumption and define what is called a *soft margin condition* - this is an implicit form of *penalisation.* Imposing these conditions allows to select a *good* candidate in the model that will hopefully have a low *generalisation error* - and this candidate will be defined by an optimisation program and what is called *Support Vector Machines or SVM.* There are three main tasks in this project.

1. Present the problem in the separable and non separable case, and introduce the margin and soft margin conditions. What can be said about the respective optimisation problems?
2. Present and prove the margin bound for binary classification (Theorem 4.4 in Mohri et.al (2012)).
3. Implement SVM for distinguishing handwritten 0 and 1.

References : The book Mohri et.al (2012), Chapter 4.

1.2.2. Regression

A problem that is quite related to classification is non-parametric regression. Here the data $\mathcal{D}_n$ are of the form $(X_i, Y_i)_i$ where the X_i are the points that are sampled in an i.i.d. fashion, and the

$$Y_i = \theta^*(X_i) + \epsilon_i,$$

are noisy evaluations of the function where ϵ_i is a i.i.d. white noise. $\mathcal{P}$ is a set of functions, and the loss function L is

$$L(\mu, \theta) = \mathbb{E}_{(X,Y)\sim\mu}(\theta(X) - Y)^2.$$

One has for the empirical loss

$$L(\mu_n, \theta) = \frac{1}{n}\sum_i (Y_i - \theta(X_i))^2.$$

Linear regression in high dimension : the lasso We are first going to consider a linear regression setting. Here the data are of the form $(X_i, Y_i)_{i\leq n}$, where

$$Y_i = \langle X_i, \beta\rangle + \epsilon_i,$$

where Y_i and ϵ_i are n-dimensional vectors (and ϵ is a noise such that $\|\epsilon\|_2 \leq \eta$), where β is the p dimensional unknown parameter, and the design X_i are p dimensional vectors.

We will assume here that $n \ll p$. Therefore, the problem is ill-posed and nothing can be done on the mode $\mathbb{R}^p$.

Question : How can we deal with this ill posed problem and penalise for finding a good model for this linear regression problem?

Project 4 : The Lasso, a specific example of regularisation

Objective : In this project, we will consider a specific example of regularisation, the lasso for. We will assume here that $n \ll p$. Therefore, the problem is ill-posed and nothing can be done on the mode $\mathbb{R}^p$: one has to restrict the model. A classical restriction is to consider models

$$\mathcal{P}(k) = \{\theta \in \mathbb{R}^p : \|\theta\|_0 \leq k\},$$

and to consider the regression loss on these nested models. On each model, one can (in theory) compute an estimator. It is assumed that the parameter θ is sparse itself, so there is one of the $\mathcal{P}(k)$ that is a good model for it with relatively small k - but one does not know k. The objectives of this project are first to understand how one can penalise in this case in order to select the right model. Then one wants to understand how one can *compute* these estimates in practice. This will be done in the following tasks.

1. Present the l_0 minimisation problem in Theorem 3.6 of Fornasier and Rauhut (2008) and the associated Null Space Property (NSP). Write the Lagrangien of the minimisation problem and explain why this is penalisation. Explain why this solution is impractical (see Fornasier and Rauhut (2008)).
2. Present the *Restricted Isometry Property (RIP)* (definition 3.2 in Fornasier and Rauhut (2008)) and present the link with the NSP. Present Theorem 3.6 in Fornasier and Rauhut (2008) (without proof) and explain what this theorem means.
3. Present Theorem 3.5 in Fornasier and Rauhut (2008) and its proof, see also Shah (2013). Write the Lagrangien of the minimisation problem and explain why this is penalisation. Explain why this solution is this time practical. The Lagrangien is the *Lasso* estimator, see Section 10.3.4 in Mohri et.al (2012).

References : The book Fornasier and Rauhut (2008) and Section 10.3.4 in Mohri et.al (2012).

Non linear regression Linear regression is the most basic example of regression, that is very common in econometrics. But quite often, this is not rich enough for representing complex data and one then wants to find more sophisticated representations for the data. The general idea of most non linear regression method is to project the data points X in a richer and larger space and perform a penalised linear regression in this new space for choosing an appropriate representation by penalisation.

Question : What are desirable properties of good representation spaces, and suitable penalties?

Project 5 : Kernel regression and RKHS

Objective : A possible choice of richer space is done through a *Kernel*, i.e. a bilinear form $K(X, X')$, and the convolutions of this linear form with the data. There are three main tasks in this project.

1. Present the concept of Kernel and RKHS (Chapter 5 in Mohri et.al (2012)).
2. Present and prove the generalisation bounds for RKHS (Chapter 10 in Mohri et.al (2012)).
3. Implement a RKHS on synthetic data.

References : The book Mohri et.al (2012), Chapter 5 and 10.

1.3. Dimension reduction

When the data are very high dimensional, it is not always very convenient to do ERM. Indeed, the computational complexity of ERM often scales with the dimension of the data and not always only with the size of $\mathcal{P}$. So it is often desirable to try to find techniques that allow to reduce the dimension of the data, in order to diminish the computational complexity. Such techniques are called *dimension reduction* techniques.

Question : How can we in general reduce the dimension of the data without losing something about them?

Project 6 : Random projections and the Johnson Lindenstrauss Lemma

Objective : Consider n points $X_1, \ldots, X_n$ that are in dimension p and assume that $p \gg n$. In some sense, given the fact that there are n only points, one does not need for most application to be in dimension p if p is much larger than n : intuitively, there should be a space of much smaller dimension than p where the n points are well represented, i.e. there should be a sub-space V of $\mathbb{R}^p$ where the projection on V of the points $X_1, \dots, X_n$ gives a good idea of the respective relations (distance) of the X_i.

In more technical terms, these spaces V should satisfy, if we note Π_V for the projection on V, and for any $(i,j) \in \{1,...,n\}^2$

$$\|X_i - X_j\|_2 \approx \|\Pi_V(X_i) - \Pi_V(X_j)\|_2.$$

A very interesting question is the following.

Question : How can we construct such spaces that are of dimension as small as possible?

Then one can project the points in a such space V and do ERM in V : since the distances in a such space are preserved, one does not loose much by projecting in it.

The objective of this project is to provide a simple construction of a good projection space V trough *random projections*.

1. Give a lower bound on the dimension of such a space for n points.
2. Prove Johnson Lindenstrauss Lemma for random projections.
3. Implement random projections and compute the distances. Use it to perform classification.

References : The paper Dagstupa (2003)

---

2. Online and Active learning

Until now, we have been considering settings where all data are available before hand. But this is not always the case and in many applications, the data become available gradually - this is the *online learning* setting. Sometimes, the learner does even have an impact on how the data is collected - this is a specific case of the online learning setting, which is called *active learning*. We are going to investigate the classical theory of active learning and a specific and simple example of active learning, which is called the *bandit problem*.

2.1. Online learning and prediction with expert advice

The classical online learning setting is also called "prediction with expert advices". The idea is the following. There are K experts and at each time t, each expert k makes a prediction $f_{k,t}$ according to some internal mechanism. At each time t, and based on the data $(X_{k,t})_k$, the learner has to make a prediction p_t. Then the environment reveals the “truth” f_t, and the learner incurs a loss

$$l_t = l(f_t, p_t).$$

The performance of the learner at time T is measured by the cumulative loss

$$L_n = \sum_{t \leq T} l_t,$$

and the objective is to make the *regret* with respect to the best expert

$$R_n = \sum_t l_t - \min_k \sum_t l(f_t, f_{k,t}),$$

as small as possible (in expectation or with high probability), with respect to the prediction $(p_t)_t$ of the learner.

Game 1: The online learning game.

```
Unknown data: $(f_{k,u})_{k,u}$
Known parameters: $K$
for $t = 1, \dots,$ do
    The experts make their predictions $(f_{k,t})_k$
    The player makes prediction $p_t$
    The environment reveals $f_t$
    The player incurs a loss $l_t$
end for
Objectives : Minimize over $(p_t)_t$ the regret $R_n$
```

Global objective : Propose good strategies for solving Game 1.

Project 7 : Prediction with expert advices

Objective : Here we make the assumption that the “truth” f_t revealed by the environment is bounded by 1, and that although the sequence $(f_t)_t$ is fixed before the start of the game, it can be arbitrary. But otherwise it can be anything, and in fact the environment can even be what is called “adversarial”, which means that since it is oblivious to the strategy of the learner for predicting the $(p_t)_t$, it can even choose the $(f_t)_t$ in a “mean” way with respect to the learner strategy. For this reason, the strategies here cannot be deterministic, in order to surprise the environment so that it cannot adapt to the player’s strategy by choosing the $(f_t)_t$ in a too “mean” way before the beginning of the game.

The objective of this project will be to present bounds on the regret of an online learning strategy, called Exponential Weights, or Weighted Majority. The tasks are the following.

1. Assume first that there is an expert that is “perfect”, i.e. there exists k such that for any t, $f_{k,t} = f_t$. Can you propose a strategy for solving this case and bounding the regret of this strategy?

2. Second, we will want to present the Exponential Weights algorithm and an upper bound on its expected regret. This can be found in the lecture notes Rakhlin (2014) (and also in Cesa-Bianchi et.al (2006) for a more extended reference). Try to provide intuition on the mechanism of the algorithm and on the proof.
3. Implement the algorithm Exponential Weights on synthetic data. Try to illustrate the bound.

References : The lecture notes Rakhlin (2014) and also the book Cesa-Bianchi et.al (2006) for a more extended reference.

2.2. The bandit problem

The bandit setting is also an online learning setting, but now the learner cannot observe all "expert advices" at the same time ; at each time t it can only observe one of the "advices".

Now let us say this in a more specific way. The learner can sample K data sources (the "experts" of before) which are often referred to as "arms". At each time t, the data source k outputs a sample $X_{k,t}$ according to some internal mechanism. At each time t, the learner does not observe the output, but can choose one of the systems $k_t \in \{1, \dots, K\}$ it wants to observe. This decision is not based on the data emitted by the system at time t (which the learner does not observe) but on the data observed in the past $(X_{k_u,u})_{u<t}$. After choosing k_t, it receives $X_{k_t,t}$. At the end of the game at time n (the game is said to be of horizon n), the performance of the learner is measured by

$$L_n = \sum_t X_{k_t,t},$$

and the objective is to make the *regret* with respect to the best arm

$$R_n = \max_k \sum_t X_{k,t} - \sum_t X_{k_t,t},$$

as small as possible (in expectation of in high probability), with respect to the arm selection $(k_t)_t$ of the learner.

Game 2: The bandit game.

```
Unknown data: $(X_{k,u})_{k,u\leq n}$
Known parameters: $K$ and $n$
for $t = 1, \dots, n$ do
    The player chooses $k_t \in \{1, \dots K\}$
    The system $k_t$ reveals the reward $X_{k_t,t}$
end for
Objectives : Minimize over $(k_t)_t$ the regret $R_n$
```

Global objective : Propose good strategies for solving Game 2.

Project 8 : The stochastic bandit problem - upper bounds

Objective : We will consider in this project the case of a *stochastic bandit*, i.e. a bandit problem where each arm k output data that are i.i.d. according to a distribution ν_k, i.e.

$$\forall k, \forall t, X_{k,t} \sim_{i.i.d.} \nu_k.$$

We will assume that all ν_k are positive and bounded by 1.

The objective of this project will be to present tight (in the sense of next project...) bounds on the regret of a bandit strategy, UCB. There are two kind of bounds that exist for bandit strategies : *problem dependent* and *problem independent* bounds. The problem dependent bounds make the parameters of the problem appear - typically the mean of the arms. The problem independent bounds do not make these problem dependent quantities appear - and are therefore valid for all bandit problems. The objective of this project will be to present these two kinds of bounds. The tasks are the following.

1. We will first want to present the problem dependent upper bound for UCB (Theorem 2.1 in the survey Cesa-Bianchi et.al (2006)). A sketch of the proof has to be highlighted.
2. Second, we will want to present the problem independent bound for UCB (Section 2.4.3 in the survey Cesa-Bianchi et.al (2006)). Again a sketch of the proof has to be presented.
3. Implement the algorithm UCB on synthetic data. Both bounds have to be illustrated on the synthetic experiments.

References : All references for this project can be found in Bubeck et.al (2012), and also in Cesa-Bianchi et.al (2006) for a broader perspective.

Project 9 : The stochastic bandit problem - problem independent lower bound

Objective : We will consider in this project the case of a *stochastic bandit*, i.e. a bandit problem where each arm k output data that are i.i.d. according to a distribution ν_k, i.e.

$$\forall k, \forall t, X_{k,t} \sim_{i.i.d.} \nu_k.$$

We will assume that all ν_k are positive and bounded by 1. Let us call $\mathcal{C}_{n,K}$ for the class of all such bandit problems.

The objective of this project will be to present what is called a *lower bound* on the regret of any bandit strategy. In other word, the objective is to find $\rho_{n,k} \geq 0$ such that for any bandit strategy of the learner, there exists always a "worst case" bandit problem in $\mathcal{C}_{n,K}$ such that the expected regret of the strategy is larger than $\rho_{n,k}$, i.e. we want to find $\rho_{n,k}$ such that

$$\inf_{bandit\ algo} \sup_{bandit\ problem\ in\ \mathcal{C}_{n,K}} R_n \geq \rho_{n,k}.$$

This approach is called minimax. The tasks are the following.

1. Present in a more precise fashion the minimax framework.
2. Present Pinsker's inequality and discuss its meaning
3. Present the problem independent lower bound for the stochastic bandit problem (Theorem 3.5 in the survey Cesa-Bianchi et.al (2006)).

References : All references for this project can be found in Bubeck et.al (2012), and also in Cesa-Bianchi et.al (2006) for a broader perspective.

Project 10 : The adversarial bandit problem

Objective : We will consider in this project the case of a *adversarial bandit*, i.e. a bandit problem where *the only assumption on the distribution is that all samples are such that* $|X_{k,t}| \leq 1$ - but note that all $(X_{k,t})_{k,t}$ are fixed before the beginning of the game by the arms. This setting is called adversarial because the sequences $(X_{k,t})_{k,t}$ can be taken in a "mean" way with respect to the learner strategy (the environment is supposed oblivious of the learner strategy). For this reason, the strategies here cannot be deterministic so that it surprises the arms.

The objective of this project will be to present bounds on the regret of an adversarial bandit strategy, EXP3. The tasks are the following.

1. Explain what is the difference between adversarial and stochastic bandit setting, and why UCB cannot be expected to provide good results here (Sections 1 and 2 in the survey Cesa-Bianchi et.al (2006)).
2. Second, we will want to present the regret bound for EXP3 (Theorem 3.1 in the survey Cesa-Bianchi et.al (2006)). A sketch of the proof has to be presented.
3. Implement the algorithm EXP3 on synthetic data, and construct synthetic data examples that are as "difficult" as possible for EXP3.

References : All references for this project can be found in Bubeck et.al (2012), and also in Cesa-Bianchi et.al (2006) for a broader perspective.

2.3. Infinitely many armed bandits

In the previous bandit problems presented, it was assumed that the number of arms K is small with respect to the time horizon n. This is not always the case in application of course, and then it is an interesting question to try to understand what it is possible to achieve. Here we will consider even *infinitely many armed bandits* (which can be a valid approximation for bandits with very many arms). We will consider a *stochastic bandit*, i.e. a bandit problem where each arm $x \in \mathcal{X}$ (where $\mathcal{X}$ is a continuous set) output data that are i.i.d. according to a distribution ν_x, i.e.

$$\forall x \in \mathcal{X}, \forall t, X_{x,t} \sim_{i.i.d.} \nu_x.$$

We will assume that all ν_k are positive and bounded by 1.

There are roughly two possibilities in this setting : either the arms are completely *unstructured* (this is the same setting as presented before, but with many arms), or there is some *structure* on the arms - and in this case, we will assume that the index x of the arms provide information on them. Depending on this, on can achieve different outcomes.

Project 11 : Infinitely many armed bandits without structure

Objective : In this project, we will assume that the arms are completely unstructured, and that we can sample randomly among them according to a probability on $\mathcal{X}$. Let μ_x be the mean of distribution ν_x, and let $\mathcal{P}$ be the distribution of μ_x when x is chosen at random according to the probability on $\mathcal{X}$. Let

$$\mu_* = \sup_{x \in \mathcal{X}} \mu_x.$$

What we will assume is that we have a lower bound on the proportion of near optimal arms, i.e.

$$\mathcal{P}(|\mu_* - \mu_x| \leq \epsilon) \geq C\epsilon^{\beta},$$

where $C > 0, \beta > 0$ are two constants.

The objective of this project is to present UCB-AIR, an algorithm for this setting, and an upper bound on its regret (see Paper Wang et.al (2008)). This will be done in two tasks.

1. Present and discuss the algorithm UCB-AIR. Discuss the parameter β. Compare UCB-AIR with UCB.

2. Present a sketch of the proof of the regret bound for UCB-AIR.
3. Implement the algorithm UCB-AIR, for various values of β.

References : A survey on this is to be found in Section 1.2.1 of the survey Munos (2014) and UB-AIR is investigated in the paper Wang et.al (2008).

Project 12 : Infinitely many armed bandits with structure - Optimisation using bandits

Objective : In this project, we will assume a *functional* structure on the arms and in fact the bandit formalism will be used to solve non-convex optimisation problems. Let μ_x be the mean of distribution ν_x. We will assume that $\mu_x = f(x)$, and that f is *s-Hölder smooth*. In this setting, aiming at minimising the regret is equivalent to aiming at solving a cumulative optimisation task - since the aim is in fact to sample as often as possible close to the optimum of f.

The objective of this project is to present the algorithm HOO that is aiming at this setting, as well as an upper bound its regret (see Section 3 of the survey Munos (2014) and the paper Bubeck et.al (2009)). This will be done in three tasks.

1. Present and discuss the hierarchical partitioning of the space and the local smoothness assumptions that are assumed for HOO, and discuss the algorithm HOO.
2. Present a sketch of the proof of the upper bound on the regret of HOO.
3. Implement the algorithm HOO and run it on some classical smooth functions. When is HOO performing well? Badly?

References : A survey on this is to be found in Section 3 of the survey Munos (2014) and in the paper Bubeck et.al (2009).

References

Boucheron, Stphane, Gbor Lugosi, and Pascal Massart. Concentration inequalities: A nonasymptotic theory of independence. *OUP Oxford*, 2013.

Bubeck, Sebastien, and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1-122, 2013.

Bubeck, S., Stoltz, G., Szepesvri, C., and Munos, R. Online optimization in X-armed bandits. *Advances in Neural Information Processing Systems*, 2009.

Cesa-Bianchi, Nicolo, and Gbor Lugosi. Prediction, learning, and games. *Cambridge University Press*, 2006.

Dagstupa Sanjoy. An elementary proof of the Johnson Lindenstrauss Lemma.

Fornasier, Massimo, and Holger Rauhut. Recovery algorithms for vector-valued data with joint sparsity constraints. SIAM Journal on Numerical Analysis 46.2, 2008.

Gyorfi, Laszlo. Principles of nonparametric learning. *Springer*, 2002.

Massart, Pascal. Concentration inequalities and model selection. *Vol. 6. Berlin: Springer*, 2007.

Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. Foundations of machine learning. *MIT press*, 2012.

Munos, Remi. From bandits to Monte-Carlo Tree Search: The optimistic principle applied to optimization and planning. *Foundations and Trends in Machine Learning*, 7(1):1-130, 2014.

Rakhlin, Alexander. Lecture notes on online learning.

Shah, Rajen. High-dimensional data and the Lasso.

A.B. Tsybakov. Introduction to nonparametric estimation. Springer Science & Business Media, 2008.

Wang, Yizao, Jean-Yves Audibert, and Rmi Munos. Algorithms for infinitely many-armed bandits. Advances in Neural Information Processing Systems, 2009.