Nowscasting GDP directional change with an application to French business survey data

Matthieu CORNEC(*)(**), Fanny MIKOL(*)

(*)INSEE
(**) CREST

Introduction

In many fields of studies, observations must be classified into two differents groups. Here are some examples of classification problems :

- predict if patient, hospitalized for a heart attack, will have a second heart attack. The classification can be based on clinical measurements.
- predict if a websurfer will buy a product on a website, knowing his income, his age,...
- guess the number of a handwritten code from a digitized image.

For their nowcasting exercise, economists have mainly focused on point forecast, rather than on directional change forecast (or sign forecast). For example, in the Note de Conjoncture [9], the GDP growth forecast for the flash estimate of Q1 2011 was 0.6%. An abundant litterature on level forecasts has flourished over the past decades (see [18, 28, 24] and references therein). For the French economy, we can quote [3, 10, 8].

Surprisingly, there are few studies examining the predictability of the sign of GDP growth movement. Methodology for sign forecast evaluation was first introduced for market timing in [17, 13] and for macroeconomic forecasts in [26, 23]. A nonparametric test of predictive performance can be found in [20]. For example, [24] use [20] test to investigate the ability of their level based models to predict GDP directional change from one quarter to another. This lack of empirical studies is all the more surprising that, at the same time, economic forecasters describe their baselined scenario in terms of acceleration or deceleration : "production accelerates", "activity slows down", "business climate remains stable". This qualitative economic analysis is mainly supported by two reasons. First, this realistic view acknowledges that nobody can predict future evolution of the economic outlook with absolute certainty (see for a detailed review on density forecast [28]). Secondly, while accuracy, as measured by quantitative errors, is important, it may be even more crucial to accurately forecast the sign of change (see [25]). Will real GDP accelerate (or decelerate) ? Economists thus underline the main direction of the economy. In other words, their review refers implictly to a sign forecast.

The purpose of this paper is threefold. First, to provide a sequential framework to analyse sign forecasts. Indeed, the forecasting exercise is made sequentially on the basis of real time data. Our framework differs from previous studies in two main points : it is model free -to allow a large class of predictors- and non asymptotic -since in practice, the number of GDP forecasts are rather limited -. The procedure to test whether forecasts are useful is based on a comparison with naive uninformed benchmarks. Second, we conduct an empirical comparative study between classification models, including linear discriminant analysis, support vector machine, regression trees and level-based models including like regression. Eventually, we construct a profile index which describes in probability terms the risk of an upcoming deceleration.

This paper is organised as follows. Section 2 describes our data sets and defines the problem set up. In section 3, we conduct a comparative empirical study for a wide range of methods from econometrics to machine learning theory. In Section 4, we derive analytical properties for future sign forecasts, namely asymptotic and non asymptotic confidence interval. Eventually, in section 5, we construct our directional risk index to describe in probability terms the uncertainty associated to our sign forecast.

1 Data description and problem set up

Flash estimate

Our goal is to forecast the directional change of the *flash estimate GDP growth rate, denoted by* y_t. Our quarterly historical data of French GDP first release starts from 1988 Q1. Flash estimates are published only 45 days after the end of the current quarter. A natural question when defining the sign change is then : should we compare the flash estimate y_t to the flash estimate y_{t-1} of the previous quarter or to its updated version ?

The two issues are of great interest, as quarterly accounts are updated at each new publication. When the results of quarter Q are first released (during quarter Q+1), the figures for quarter Q-1 are updated and they are then likely to slightly differ from the ones released on quarter Q. Instead of forecasting the evolution between the two first releases of GDP growth for quarters Q-1 and Q, it would be then interesting to predict the evolution between results of Q-1 and Q that will both be published in Q+1. Table 1 illustrates these two conventions, which can lead to different forecasts for GDP growth sign.

Table 1 – Example of successive GDP growth rate publications

	Date of release	
	Q	Q+1
Reference quarter		
Q-1	0,3%	0,5%
Q	-	0,4%

Suppose that during forecasting exercise in quarter Q, we predict a growth acceleration between Q-1 and Q. This means, in the first approach, that we forecast that the first release of quarter Q growth will be greater than the available first release of Q-1 growth $(0,3\%)$. In this approach, our forecast will be correct since first preliminary figures of Q growth, published in Q+1, is effectively higher $(0,4\%)$. In the second approach though, an acceleration forecast means that we predict the first release of Q growth, published during Q+1, to be higher than the *updated* value of Q-1 growth published on the same date $(0,5\%)$. In this second approach, we then fail to predict GDP directional change correctly with this example. It is thus crucial to precisely define the starting point of our sign forecasts. The choice made in this study to focus on the first approach (i.e. with the example above, comparing $0,4\%$ with $0,3\%$) is justified as follows : during the forecasting exercise when we only know preliminary figures of Q-1, comments made about point forecast for quarter Q implicitely compare the latter with the first release of GDP Q-1 growth, no matter how this Q-1 growth is going to be revised in the future.

Business surveys

In this study, the economic information included in x_q, apart from past GDP observations, will be business surveys. Indeed, they are a useful source of information when forecasting, for they present three types of advantages : (1) they provide reliable information coming directly from the economic decision makers, (2) they are rapidly available (about a month after the questionnaires are sent), on a monthly, bimonthly or quarterly basis, and (3) they are subject to small revisions (each publication presents a generally negligible revision, only on the preceding point). Insee surveys among business leaders of all sectors are monthly qualitative surveys providing information on the rate of activity in the recent past, during the current month and in the near future. Industry survey for example questions 4,000 entrepreneurs about recent and probable future trends in their production, about their total and foreign order-book levels, inventory levels and general output prospects. Generally, these questions are qualitative : up, no change or down. The balance of opinion, defined as the difference between the percentage of positive responses and the percentage of negative responses, is the most widely-used indicator by outlook analysts to summarise answers to a question. Insee also publishes a composite indicator called the french business climate indicator : it summarises information that is common to a set of 26 balances of opinion of all sectors business surveys [6]. Insee business surveys provide the best advanced indicator for the output of the current quarter. Indeed, during our forecast exercise at the middle of quarter Q, results of business surveys for the two first months of quarter Q are available. Quantitative indicators for month M (such as industrial production index,...) are most of the time not available before the end of month $M+1$: they won't be taken into consideration here.

Sequential framework

In a sequential version of the sign prediction problem, the economic forecaster is asked to guess the next direction of the flash *estimate* of the quaterly GDP growth rate at quarter q denoted y_q. We define *the GDP directional change* between quarters $q-1$ and q as :

$$\varepsilon_q := 1\{y_q \geq y_{q-1}\}.$$

Sign serie (ε_q) for the period [2000Q1 ;2010Q4] is for example shown below :

With this notation, we say that there is an *acceleration* at quarter q (respectively *deceleration*) if $\varepsilon_q = 1$ (resp. $\varepsilon_q = 0$).

Following the sequential framework in [4], at each quarter $q = 1...Q$, we observe the economic data $x_q \in \mathbb{R}^d$, containing various economic information, including past GDP observations. At quarter q though, the first release y_q and thus the sign change ε_q are unknown. Indeed, for Insee "Conjoncture in France" [9], economists forecast GDP growth of the current quarter during the middle of the quarter. However, national accounts will

	Qtr1	Qtr2	Qtr3	Qtr4
2000	0	0	0	1
2001	0	0	1	0
2002	1	1	0	1
2003	1	0	1	1
2004	1	1	0	1
2005	0	0	1	0
2006	1	1	0	1
2007	0	0	1	0
2008	1	0	1	0
2009	1	1	1	1
2010	0	1	0	0

publish the flash of the current quarter (quarter q) during the middle of the next quarter (quarter $q+1$).

The economic forecaster is thus asked to guess the next outcome ε_q of a sequence of binary outcomes ($\{0,1\}$) $\varepsilon_1, ..., \varepsilon_{q-1}$with the knowledge of the past $\varepsilon_{\overline{q-1}} := (\varepsilon_1, ..., \varepsilon_{q-1})$ and the side economic information $x_{\bar{q}} := (x_1, ..., x_q)$: is there an acceleration or a deceleration ? In other words, the elements $\varepsilon_1, \varepsilon_1, ...$ and $x_1, x_2, ...$are revealed one at a time, beginning with $(x_1, y_0), (x_2, y_1)$...and the forecast at quarter q is based on $\varepsilon_{\overline{q-1}}$ and $x_{\bar{q}}$.

To evaluate a forecasting strategy, we need then to simulate it on a realtime basis. This appears possible thanks to the nature of our data. Recall indeed that flash estimates are the first releases of GDP growth rates as published by national accounts and thus are realtime data. Business surveys, at last, can be considered as "pseudo-realtime" data. Indeed, balances of opinion are only slightly revised over time. French climate indicator, though, was first published in January 2009, and its coefficients are re-estimated every year, possibly inducing slight revisions. However, this synthetic indicator appears to be very stable over time [2], and can then also be considered as "pseudo-realtime" data.

Formally, a *forecasting strategy* is defined as a family of predictors $(\phi_q)_q$:

$$\widehat{\varepsilon_q} := \phi_q(x_{\overline{q}}, \varepsilon_{\overline{q-1}}).$$

Loss function

To evaluate a forecasting strategy, we need to define a loss function. At quarter q, *the (normalized) cumulative loss* related to the strategy (ϕ_q) is defined as :

$$L_q(\phi_q) := \frac{1}{q} \sum_{t=1}^{q} 1(\varepsilon_t \neq \hat{\varepsilon}_t).$$

This loss function is equivalent to the average error rate $\eta_t := 1\{\hat{\varepsilon}_t \neq \varepsilon_t\}$ over the period $[1, q]$. A natural loss function for risk adverse agents could make misclassified decelerations more costly than misclassified accelerations. However for sake of simplicity, we restrict our analysis to the symmetric indicator loss. In the next sections, we compare forecasting errors for different strategies since 1997, date of the first available quarterly Insee GDP forecast.

At this point, obviously, the mean error of any strategy takes its value in $[0, 1]$. $L_q(\phi_{\overline{q}}) = 0$ (respectively $L_q(\phi_{\overline{q}}) = 1$) means that the strategy ϕ_q perfectly forecasts the sequence of signs (resp. was wrong at all time). But in between, what is a good strategy ? In this study, our approach is to compare any strategy with the optimal forecasting strategy of an uninformed forecaster trying to minimize his loss. To minimize in probability the worst case over all possible outcomes, the uninformed agent forecast the next sign change by simply drawing a random variable from a Bernouilli with parameter 1/2.

$$\widehat{\varepsilon_q} = \phi_q^{random}(\varepsilon_{\overline{q-1}}) = u_q.$$

A simple model free test

The expected normalised loss is then $\mathbb{E}L_q(\phi_q^{random}) = 1/2$, which means that on the long run, the average error of a uninformed forecaster will be equal to 0.5. Recall that $qL_q(\phi_q)$ is distributed as a binomial distribution with parameters $n = q$ and $p = 1/2$. Thus, with great probability (95%) over the period [1997Q1 ;2010Q4] (52 quarters), the mean error is greater than 35%. A strategy is all the better than it is unlikely that an uninformed forecast reaches the same error rate. We define the p-value of a strategy ϕ_q as

$$p_q := \mathbb{P}(L_q(\phi_q^{random}) \leq L_q(\phi_q)).$$

By convention, we say that a strategy is significant at level α if $p_q \leq \alpha$. In other words, if $p_q \leq \alpha$, it means that a random forecast has less than α% chances to get an error rate lower than the error rate $L_q(\phi_q)$ of the strategy ϕ_q.

2 Empirical comparison on historical data

In this section, we conduct a comparative study between different class of models ranging from econometrics to machine learning.

2.1 Naive strategies

Freeze strategy

Notice that, since 1997, quaterly signs (ε_q) have alternate about 63% of the time. A strategy would be to predict, for each quarter, the opposite sign of the previous observed direction change :

$$\widehat{\varepsilon_q} := \phi_q^{prev.value}(\varepsilon_{\overline{q-1}}) = 1 - \varepsilon_{q-1}.$$

Numerical application : Freeze mean error over the period [1997Q1 ;2010Q4] is 36%. The corresponding p-value is $1,8 \times 10^{-2}$.

Long term mean forecast

This strategy consists in assuming the current quarter sign to be equal to the long term mean of the past observed signs (which is then rounded to 0 or 1).

$$\widehat{\varepsilon_q} = \phi_q^{longterm}(\varepsilon_{\overline{q-1}}) = 1\left\{\frac{1}{q}\sum_{t=1}^{q}\varepsilon_t \geq 0,5\right\}.$$

Numerical application : Long term strategy mean error is 41%. The corresponding p-value is $5,0 \times 10^{-2}$.

Markov strategy

Here we assume $\varepsilon_{\overline{q}}$ to be the outcome of a homogeneous Markov chain of order 1 with probability transition matrix Π. Forecast is then given by :

$$\widehat{\varepsilon_q} = \phi_q(\varepsilon_{\overline{q-1}}) := Argmax_{i \in \{0,1\}} \hat{\Pi}(\varepsilon_{q-1}, i)$$

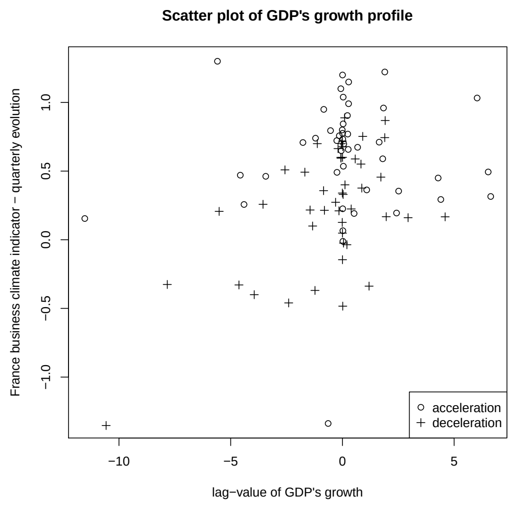
where $\hat{\Pi}(\varepsilon_{q-1}, i) := \hat{\mathbb{P}}(E_q = i | E_{q-1} = \varepsilon_{q-1})$.

In other words, suppose we observed at $q-1$ that $\varepsilon_{q-1} = j$, then we will forecast for quarter q the most likely occurrence of E_q when its previous value was j.

Numerical application : Markov strategy mean error since 1997 is 36%. The corresponding p-value is $1,8 \times 10^{-2}$.

2.2 Business surveys based strategies

In this section, we incorporate business surveys in our strategies. Balances of opinion from business surveys provide indeed an adequate predictor of the GDP sign change. For example, figure below corresponds to GDP growth signs, with the lag value of GDP's growth in the x-axis and the French business climate growth in the y-axis. These two variables appear to be not-so-bad input variables for sign prediction : activity is obviously more likely to accelerate when business climate is accelerating and when GDP growth during the previous quarter was low. This illustrates the opportunity to test parametric classification models with business surveys as explanatory variables.



Recall the forecaster is asked to guess the next directional change $\widehat{\varepsilon_q}$ with knowledge of the past observations x_q. Thus, mean errors of the strategies presented below are computed on an out-of-sample basis.

Note : parametric methods presented in this section are computed thanks to $R^{©}$ software packages "dynlm", "MASS", "rpart", "svmpath".

Regression model based forecast

Here we derive a profile forecast from a quantitative GDP growth forecast. Quantitative forecast is obtained through an usual least-square regression model. Direction forecast is then simply obtained by comparing the point estimate of the current quarter $\widehat{y_q}$ with the observed growth rate of the previous GDP release y_{q-1}. This approach is the most commonly used (see introduction). This strategy is then defined as :

$$\widehat{\varepsilon_q} = \phi_q^{reg}(x_{\overline{q}}) = 1\left\{\widehat{y_q} \geq y_{q-1}\right\}.$$

The "core model" used to predict y_q includes the lag-value of flash estimate y_{q-1}, the level of French business climate indicator F_q for quarter q (in fact, the sum of the three previous monthly indicateur observed, see notations below) and the "signed" acceleration of this indicator $\Delta F_q \left|\Delta F_q\right|$:

$$\widehat{y_q^{PR}} = \widehat{\beta}_0 + \widehat{\beta}_1 y_{q-1}^{PR} + \widehat{\beta}_2 F_q + \widehat{\beta}_3 \Delta F_q \left|\Delta F_q\right|.$$

where $F_q = facfr_{q,1} + facfr_{q,2} + facfr_{q-1,3}$, $facfr_{q,i}$ denoting the monthly business climate indicator for the i^{th} month of quarter q. The vector $\widehat{\beta}$ stands for the ordinary least square estimates. Recall that during the forecasting exercise around the middle of quarter q, the most recent availabe indicator is $facfr_{q,2}$.

Numerical application : The normalised loss of the regression based strategy $\widehat{\varepsilon_q} = 1\left\{\widehat{y_q} \geq y_{q-1}\right\}$ is 18% since 1997. The corresponding p-value is $8,4 \times 10^{-7}$.

We also estimate a second model, which includes balances of opinion in manufacturing industry dealing with recent changes in output ($manuf.tppa_{q,i}$) and personal production expectations ($manuf.tppre_{q,i}$). Output in manufacturing industry is indeed considered as one of the best leading input variables for GDP growth :

$$\begin{array}{c} y_q = \hat{\beta}_0 + \hat{\beta}_1 y_{q-1} + \hat{\beta}_2 y_{q-4} + \hat{\beta}_3 manuf.tppa_{q,2} + \hat{\beta}_4 (manuf.tppa_{q,2} - manuf.tppa_{q,1}) \\ +\hat{\beta}_5 (manuf.tppre_{q,2} - manuf.tppre_{q,1}) + \hat{\beta}_6 \left(manuf.tppre_{q,1} - manuf.tppre_{q-1,3}\right) \end{array}.$$

Numerical application : Sign mean error with the profile based on the "manufacturing model" is 18%. The corresponding p-value is $8,4 \times 10^{-7}$.

Probit forecast

Signs ε_q can be directly predicted through a parametric probit model, with a set of explanatory variables x_q. This model states that :

$$P(\varepsilon_q = 1 | X_q = x_q) = F(\beta x_q)$$

where F is the cumulative normal distribution. Coefficients β are estimated by likelihood maximisation. Sign forecasts are then given by :

$$\widehat{\varepsilon_q} = \phi_q^{prob}(x_{\overline{q}}, \varepsilon_{\overline{q-1}}) = 1\left\{F(\hat{\beta} x_q) \geq 0,5\right\}.$$

Here we also consider two sets of variables for x_q :

- the "core" variables used in the "core model" of the previous section

$$(y_{q-1}, \Delta F_q, \Delta F_q \left|\Delta F_q\right|).$$

- the "manufacturing" variables used in the "manufacturing model" of the previous section

$$\left(\begin{array}{c} y_{q-1}, y_{q-4}, manuf.tppa_{q,2}, manuf.tppa_{q,2} - manuf.tppa_{q,1}, \\ manuf.tppre_{q,2} - manuf.tppre_{q,1}, manuf.tppre_{q,1} - manuf.tppre_{q-1,3} \end{array}\right)$$

Numerical application : The mean loss for "core" model (resp. manufacturing model) is 14% (resp. 16%) since 1997. The corresponding p-value is $3,6 \times 10^{-8}$.

"Linear/Quadratic discriminant analysis" (LDA/QDA)

Linear discriminant analysis (LDA) is a classification method. It provides linear decision boundaries depending on the observed variables X_q (for more details see [12] and references therein). This method requires the knowledge of the class posteriors $Pr(\varepsilon/X = .)$. Let $f_k(.)$ be the posterior density of X given class $\varepsilon = k$ ($k \in (0,1)$), and let π_k be the prior probability of class k, with $\sum_{k=0}^{1} \pi_k = 1$. A simple application of Bayes theorem gives us :

$$Pr(\varepsilon = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{l=0}^{1} f_l(x)\pi_l}.$$

Suppose each class density $f_k(x)$ is a multivariate Gaussian with mean μ_k and covariance matrix Σ_k. Linear discriminant analysis arises in the special case when we assume that the classes have a common covariance matrix $\Sigma_k = \Sigma \ \forall k$ (otherwise we apply the term "quadratic discriminant analysis" - QDA - where more factors are to be estimated). In practice, the parameters (π_k, μ_k, Σ_k) are unknown, so we must estimate them through our historical data :

- $\pi_k = N_k/N$, where N_k is the number of class-k observations,
- $\widehat{\mu_k} = \sum_{\varepsilon_q = k} x_q/N_k$;
- $\widehat{\Sigma} = \sum_{\varepsilon_q = k}(x_q - \widehat{\mu}_k)(x_q - \widehat{\mu_k})^T/(N-K)$.

The LDA rules can be written as a function of these estimated parameters. More precisely, the LDA rule classifies the observations x_q in class $k = 1$ rather than class $k = 0$ if :

$Pr(\varepsilon_q = 1|X_q = x_q) > Pr(\varepsilon_q = 0|X_q = x_q)$

$\iff x^T\Sigma^{-1}(\widehat{\mu_1} - \widehat{\mu_0}) > \frac{1}{2}\widehat{\mu_1}^T\widehat{\Sigma}^{-1}\widehat{\mu_1} - \frac{1}{2}\widehat{\mu_0}^T\widehat{\Sigma}^{-1}\widehat{\mu_0} + log(N_0/N) - log(N_1/N).$

We see that decision boundaries correspond to linear functions of the explanatory variables x_q. Profile forecast is then given by :

$$\widehat{\varepsilon_q} = \phi_q^{LDA}(x_{\overline{q}}) = 1\left\{x_q^T\Sigma^{-1}(\widehat{\mu_1} - \widehat{\mu_0}) > \frac{1}{2}\widehat{\mu_1}^T\widehat{\Sigma}^{-1}\widehat{\mu_1} - \frac{1}{2}\widehat{\mu_0}^T\widehat{\Sigma}^{-1}\widehat{\mu_0} + log(N_0/N) - log(N_1/N)\right\}.$$

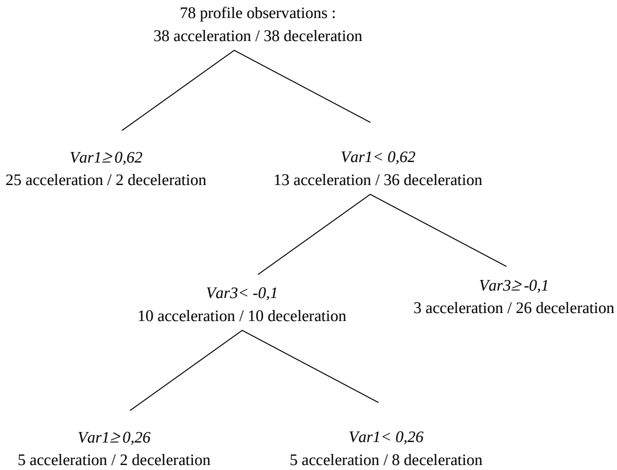
Notice that with two classes there is a simple correspondence between LDA and classification by linear least squares. This is equivalent to directly estimate the model $\delta_q = \beta_0 + \beta_1 y_{q-1}^{PR} + \beta_2 F_q + \beta_3 \Delta F_q \left|\Delta F_q\right| + \xi_q$, then to classify $\widehat{\varepsilon_q}$ according to the position of $\widehat{\delta_q}$ compared to 0.5.

Numerical application : We consider for X_q the "core variables" presented above. Mean of forecasting errors is 12% since 1997. The corresponding p-value is $6,5 \times 10^{-9}$.

We also considered the LDA-extension method, i.e. the quadratic discriminant analysis (QDA). The mean loss is a bit higher than with the LDA method (14% since 1997). Indeed, QDA implies the estimation of a lot more parameters (the variance parameters, see above) than LDA, with a limited number of observations (around 76 quarterly observations in our case). Consequently, due to an increasing complexity, results are then less accurate.

Recursing partitioning (RPART)

Classification and regression trees (as described in [5]) can be generated through the "RPART" algorithm. The goal is to predict directional change (acceleration or deceleration), on the basis of our observed economic variables x_q. The tree is built by the following process : first the single variable is found which best splits the data into two groups. The data is split, and then this process is applied separately to each sub-group, and so on until the subgroups either reach a minimum size or until no improvement can be made. The second stage of the procedure consists of using cross-validation to trim back the full tree. We use as external observations the "core" explanatory variables $(y_{q-1}, \Delta F_q, \Delta F_q\,|\Delta F_q|)$. The resultant model separated the observations into four groups as shown in figure below, where $var1 = y_{q-1}$ and $var3 = \Delta F_q\,|\Delta F_q|$. We see that the second variable of our explanatory observations set is not used to build the tree, which could explain the poor results obtained compared to other classification strategies (see below). We tried other sets of explanatory variables, but no other specification improves the performance.



Sign forecast $\widehat{\varepsilon_q} = \phi_q^{rpart}(x_{\overline{q}})$ is given by one of the final knot reached by the observation $X_q = x_q$.

Numerical application : The mean error since 1997 is 25%. The corresponding p-value is $9,1 \times 10^{-5}$.

Support vector machine (SVM)

Support vector machines (SVM, see [12]) is a generalization of linear decision boundaries for classification like LDA : SVM produces nonlinear boundaries by constructing a linear boundary in a large, transformed version of the feature space. Besides this method deals with non-separable cases, as it allows for some points to be on the wrong side of the margin. These misclassified points are penalized with a certain cost parameter C which has to be set. A high C means for example that we want to minimize the number of misclassified datapoints : notice that this can lead to an overfitted model which fits well training data but do wrong for forecast. The optimal value for C can be estimated by cross-validation.

SVM decision boundary is written as :

$$\left\{x_q \text{ such as } f(x_q) := h(x_q)^T\beta + \beta_0 = 0\right\}$$

where h can be a non-linear function of the observations x_q. The introduction of such function allows to transform the initial observations x_q space to an enlarged space $f(x_q)$: this is likely to achieve better training-class separation. This decision boundary then translates to nonlinear boundaries in the original space. Next, parameters (β, β_0) are estimated in order to optimize the decision boundary between the two classes of ε (0 or 1) :

$$min \|\beta\| \text{ subject to } \begin{cases} \varepsilon_q f(x_q) \geq 1 - \xi_q, \forall q \\ \xi_q \geq 0, \sum \xi_q \leq C \end{cases} .$$

where ξ_q denote the slack variables authorizing missclassified points (here, missclassifications occur when $\xi_q > 1$). It can be shown that the solution only involves $h(x)$ through its inner product $K(x, x') = \langle h(x), h(x') \rangle$. SVM algorithm provided by $R^{©}$ "svmpath" package allows us to choose between K kernel-functions like :

$$\begin{aligned} d\text{th-degree polynomial} &: K(x, x') = (1 + \langle x, x' \rangle)^d \\ \text{radial basis of order } \gamma &: K(x, x') = exp(-\gamma \|x - x'\|^2) \end{aligned}$$

Forecast is finally given by the classification rule induced by $f(x)$:

$$\widehat{\varepsilon_q} = \phi_q(x_{\overline{q}}) = sign\left(\widehat{f(x_q)}\right).$$

Here we take for x_q the core variables exposed above (i.e. with french business indicator's level and acceleration).

Out-of-sample forecast errors are fewer when we set a high cost parameter C (i.e. when we are close to a separable case). Hence we see here that there is no gain in enlarging our observation space with a non-linear function h. Besides, kernel-function K that leads to the best out-of-sample performance is the 1th-degree polynomial function : this is equivalent to a linear decision boundary.

Numerical application :The error rate is then 16% since 1997. The corresponding p-value is $1,8 \times 10^{-7}$.

Summary of the strategies forecasting performances

On our historical data, we summarize previous empirical performances in the table below. The two main conclusions are :

1. Using business surveys, quantitative model based strategies can slightly surpass Insee performance. Best methods include Linear Discrimant Analysis and Probit, with a error rate of 12%.

Strategy	**Error since 1997**	**p-value**
Random forecast	0,50	-
Opposite of the previous value	0,36	$1,8 \times 10^{-2}$
Opposite of the long-term mean	0,39	$5,0 \times 10^{-2}$
Markov forecast	0,36	$1,8 \times 10^{-2}$
Regression model (core variables)	0,18	$8,4 \times 10^{-7}$
Regression model (manufacturing variables)	0,18	$8,4 \times 10^{-7}$
Probit (core variables)	0,14	$3,6 \times 10^{-8}$
Probit (manufacturing variables)	0,16	$1,8 \times 10^{-7}$
LDA (core variables)	0,12	$6,5 \times 10^{-9}$
QDA (core variables)	0,14	$3,6 \times 10^{-8}$
RPART (core variables)	0,25	$9,1 \times 10^{-5}$
SVM (core variables)	0,16	$1,8 \times 10^{-7}$

So far, we have conducted an empirical comparative study on our empirical data. In other words, we have answered the question : what would have been the error rate of a forecast using such strategy from 1997 Q1 to 2010Q4 ?

A natural question is now : what can be said about future nowcasts ? What about the error rate of a chosen strategy during the next Q quarters ? To answer those questions, we must take into account different kinds of uncertainty to derive analytical properties. We tackle these issues in the next section.

3 Prediction for future sign forecasts

3.1 Test of independence

Recall the forecast errors are denoted by $\eta_t := 1\{\varepsilon_t \neq \hat{\varepsilon_t}\}$. To model uncertainty, we consider (η_q) as the outcome of a random process (H_q). In order to derive analytical properties, a central point is then to test whether forecast errors are independent or not. For example, if forecast errors of a given strategy are concentrated over the end of the period, it will raise serious doubts about the ability of this strategy to predict next GDP growth signs. Non stationary underlying processes could for instance explain such time-correlated forecast errors. For the sake of simplicity, we suppose that the time dependency can be of order 1 at the most. Recall that independence is equivalent to : for any t, i, j,

$$\mathbb{P}(H_t = i, H_{t-1} = j) = \mathbb{P}(H_t = i)\mathbb{P}(H_{t-1} = j)$$

which can be rewritten as :

$$\begin{cases} \mathbb{P}(H_t = 1, H_{t-1} = 1) & = \mathbb{P}(H_t = 1)\mathbb{P}(H_{t-1} = 1) \\ \mathbb{P}(H_t = 1, H_{t-1} = 0) & = \mathbb{P}(H_t = 1)\mathbb{P}(H_{t-1} = 0) \\ \mathbb{P}(H_t = 0, H_{t-1} = 1) & = \mathbb{P}(H_t = 0)\mathbb{P}(H_{t-1} = 1) \\ \mathbb{P}(H_t = 0, H_{t-1} = 0) & = \mathbb{P}(H_t = 0)\mathbb{P}(H_{t-1} = 0) \end{cases}.$$

Since (η_t) is a dummy function, previous equations are equivalent to test the single equation $Cov(H_t, H_{t-1}) = 0$.

Proposition 1. *With the previous asumptions, an asymptotic α-level test for the null hypothesis $H_0 := \{Cov(H_t, H_{t-1}) = 0\}$ can be defined by the following reject region :*

$$\left\{ \left| \frac{\sqrt{Q}\widehat{Cov}(H_t, H_{t-1})}{\sqrt{\hat{\lambda}'\hat{\Sigma}\hat{\lambda}}} \right| \geqslant q_{1-\alpha/2} \right\}.$$

with $\hat{\lambda} := (1, \hat{E}(Z_t^3), \hat{E}(Z_t^2))'$, $Z_i := \begin{pmatrix} H_{i-1}H_i \\ H_i \\ H_{i-1} \end{pmatrix}$, *and* $\hat{\Sigma} := \widehat{Var}(Z_1) + 2\left[\widehat{cov}(Z_1, Z_2) + \widehat{cov}(Z_1, Z_3)\right]$ *and* $q_{1-\alpha/2}$ *the* $1 - \alpha/2$ *normal quantile.*

Proof : see appendices.

Numerical application : For all the strategies presented above, we do not reject the null hypothesis of independent forecast errors. Indeed relevant tests all belong to $[-1.2, 0.5]$, that is inside the 95% confidence interval $[-1.96, 1.96]$. As a conclusion, we will assume afterwards that forecast errors of all strategies are time-independent. However, two different strategies can still have correlated forecast errors.

3.2 What is the future error rate of a given strategy ?

Asymptotic approach

For a given strategy ϕ_1, we estimated the mean of forecast errors with the historical data by $\widehat{p_1}$. We now want to predict the mean of the *future* forecast errors during the next h quarters $\eta_q, q = Q+1...Q+h$. The mean of the future forecast errors during the next h quarters will be $\frac{1}{h}\sum_{t=Q+1}^{Q+h}\eta_t$.

If h is large enough, we apply the central limit theorem and we have asymptotically with probability $1-\alpha$:

$$\frac{1}{h}\sum_{t=Q+1}^{Q+h} H_t \in p_1 \pm \frac{\sigma q_{1-\alpha/2}}{\sqrt{h}}$$

with p_1 the (unobserved) true expected error rate for strategy ϕ_1, σ is the (unobserved) stantard-deviation of η_t, and $q_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution. Thus, the length of the confidence interval is controled by a *forescast error* which converges to 0 at rate $1/\sqrt{h}$. However, since p_1 is unknown, we must estimate it through our historical dataset. Denoting Q the size of our historical sample, we obtain for large Q with probability $1-\alpha$:

$$\widehat{p_1} \in p_1 \pm \frac{\sigma q_{1-\alpha/2}}{\sqrt{Q}}.$$

In this equation, the length of the confidence interval is controlled by an *estimation error* term which tends to 0 at rate $1/\sqrt{Q}$. Combining both equations, and replacing σ by its empirical counterpart $\hat{\sigma} := \sqrt{\frac{1}{Q-1}\sum_{q=1}^{Q}(H_q-\bar{H})^2} = \sqrt{\widehat{p_1}(1-\widehat{p_1})}$, we apply the central limit theorem for large Q and h.

Proposition 2. *Under previous asumptions, we get asymptotically with probability* $1-\alpha$ *:*

$$\frac{1}{Q}\sum_{t=1}^{Q} H_t \in \widehat{p_1} \pm \hat{\sigma} q_{1-\alpha/4}\left[\underbrace{\frac{1}{\sqrt{h}}}_{Forecast\ error} + \underbrace{\frac{1}{\sqrt{Q}}}_{Estimation\ error}\right].$$

Thus, uncertainty, measured as the length of the confidence interval, results from two sources of uncertainty : forecast uncertainty and estimation uncertainty.

Numerical application : with $\widehat{p_1} \approx 0,15$ (e.g. a value close to our best strategies), we obtain :

- with $N = 60$ (15 years) and $Q = 8$ (2 years) the 90% confidence interval for the future mean of forecasts errors will be $[0, 0.5]$.
- with $N = 60$ (15 years) and $Q = 60$ (15 years) the 90% confidence interval for the future mean of forecasts errors will be$[0, 0.33]$.

These confidence intervals for future forecasts may look quite large. But notice that any of our parametric strategies is a lot more successful than the random strategy. The upper bound of the relevant 90% confidence interval for best parametric strategies (e.g. around 0.5 with $N = 60$ and $Q = 8$) is indeed around 0.5 with $N = 60$ and $Q = 8$. At the same time, an uninformed agent will only be able to give the following confidence interval for the mean of its future forecast errors : $0.5 \pm 0.5 * 2 * (1/\sqrt{60} + 1/\sqrt{8}) = [0.1]$. The corresponding upper bound hence reaches in that case the maximum possible value of 1 !

Finite sample approach

Previous results are only valid when Q and h are large (application of the central limit theorem). However, in practice, we only have access to a limited number of observations. Thus the relevance of previous bounds can be challenged. In this paragraph, we derive finite sample results to deal with this issue.

A first approach consists in using Hoeffding's inequality ([14]). The latter upperbounds the probability that the distance between empirical mean and expectation is large.

Proposition 3. *Then, with a probability lower than* $1-\alpha$*, we have :*

$$|\frac{1}{Q}\sum_q H_q - \widehat{p_1}| \leq \sqrt{\frac{\ln(4/\alpha)}{2}}(\frac{1}{\sqrt{Q}} + \frac{1}{\sqrt{N}}).$$

Proof : see appendices.

Numerical application : for $N = 60$ and $Q = 8$ the interval length around $\widehat{p_1}$ is $\pm 0,6$! ($\pm 0,3$ when $N = Q = 60$).

We lose a precision factor of order 3 in these bounds in comparison with the asymptotic approach. If we compare the two expressions, the variance term σ is missing in the finite sample bounds. Thus it is a uniform bound that does not take into account the fact that the variance can be small. To fill the gap, using inequality in [16], we derive an upper bound with an empirical variance term.

Proposition 4. *Under previous assumptions, we have with probability* $1-\delta$ *:*

$$|\frac{1}{Q}\sum_q H_q - \widehat{p_1}| \leq \underbrace{\hat{\sigma}_N\sqrt{2\ln(8/\delta)}\left(\frac{1}{\sqrt{Q}} + \frac{1}{\sqrt{N}}\right)}_{Asymptotic\ bound} + \underbrace{\frac{7\ln(8/\delta)}{3(N-1)} + \frac{\ln(8/\delta)}{3Q} + \frac{2\ln(8/\delta)}{\sqrt{Q(N-1)}}}_{Approximation\ error}.$$

Proof : see appendices.

The confidence interval length is of order $\pm 0,5$ ($h = Q = 60$). This disappointing result in comparison with the previous inequality is due to the fact that, in our case, the empirical variance $\hat{\sigma}_N$ is not small enough to compensate the approximation term. However it still might be interesting for other applications. As expected, the length of the confidence interval is larger in the finite sample approach than in the asymptotic framework.

Thus, for practical purposes, economic forecasters may consider those intervals too large. To deal with this issue, notice that the error rate can be seen as an average error over all possible outcomes x_q. In probability terms, the empirical error rate estimates the unconditionnal expected error i.e. $\mathbb{E}_{X_q,\varepsilon_q}(\varepsilon_q)$. For a particular quarter q, it is all the more interesting for the forecaster to give a conditionnal directional scenario, i.e. $\mathbb{E}(\varepsilon_q|X_q = x_q)$. Indeed, if on average, the error rate is equal to 12%, the conditional error $\mathbb{E}(\varepsilon_q|X_q = x_q)$ could be even smaller. This partial conclusion advocates for a conditional approach.

4 Directional risk index

In this section, we define a "directional risk index", which will give for each quarter the conditional probability of success (or failure) associated with our directional forecast. More precisely, it will give the probability to be in an acceleration state (e.g. $\varepsilon_q = 1$) or in a deceleration state (e.g. $\varepsilon_q = 0$). For symmetry reason, this "profile index" ranges from -1 (deceleration) to $+1$ (acceleration). We define the directional risk index by :

$$I_q := 2(\hat{\mathbb{P}}(\varepsilon_q = +1|x_{\bar{q}}) - 1/2)$$

with $\hat{\mathbb{P}}(\varepsilon_q = +1|x_{\bar{q}})$ the estimated conditional probability of being in an acceleration state given the knowledge of the business surveys up to quarter q. We can also define an "area of uncertainty" when, by convention, the directional risk index lies in the interval $[-0.5, +0.5]$. A profile forecast associated with a directional risk index that falls into the area of uncertainty has then to be considered carefully. To build this probability index, we consider two previous strategies : the regression model strategy, and the probit strategy.

4.1 Directional risk index derived from the regression method

To derive such an index, we need somehow to introduce probability in the previous strategy. In the regression approach, it is usual to assume a linear statistical model with *i.i.d.* errors such that :

$$y_q = \beta_0 + \beta_1 y_{q-1} + \beta_2 F_q + \beta_3 \Delta F_q \left|\Delta F_q\right| + \xi_q$$

This model gives for each quarter q the growth level forecast $\hat{y}_q$. Recall sign forecast was then given by :

$$\phi_q(x_{\bar{q}}) = 1\left\{\widehat{y_q} \geq y_{q-1}\right\}$$

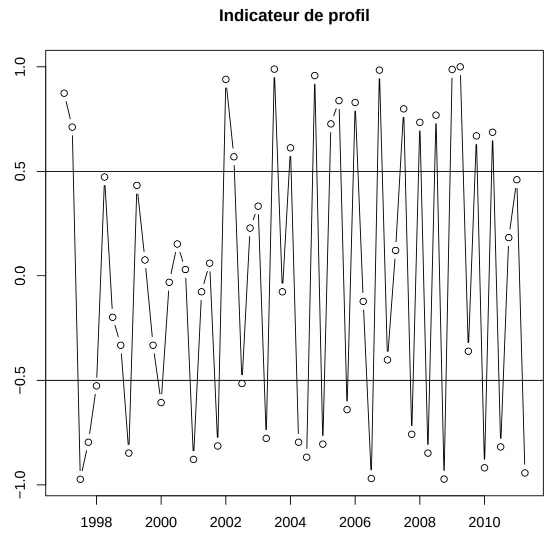
with $\widehat{y_q} = x_q^{'}\widehat{\beta}$ and $\widehat{\beta}$ the ordinary least square estimate. We estimate the density of the error term ξ_q with a kernel-type estimation :

$$\widehat{f_\xi}(x) = \frac{1}{Nh}\sum_{q=1}^{Q} K(\frac{x - x_q}{h})$$

where h is the optimal kernel bandwidth. It follows that :

$$\mathbb{P}_q(\varepsilon_q = 1) = \mathbb{P}_q(\xi_q \geq y_{q-1} - x_q^{'}\beta) \approx \mathbb{P}_{\widehat{f_\xi}}(\xi_q \geq y_{q-1} - x_q^{'}\hat{\beta})$$

The following graph provides the corresponding directional risk index $I_q := 2(\hat{\mathbb{P}}(\varepsilon_q = +1|x_{\bar{q}}) - 1/2)$ (horizontal lines delimit the area of uncertainty) :



Over our historical sample, the directional risk index falled in the "reliable area" (e.g. out of the "area of uncertainty") around 66% of the time. Inside that area, the average error rate $\{\eta_q = 1\}$ obtained through any of our best strategies (LDA, probit, regression...) falls below 4% (that is, only 2 wrong forecasts in this area over the period 1997-2010).

4.2 Directional risk index derived from the probit strategy

Recall the assumptions behind our probit forecast strategy :

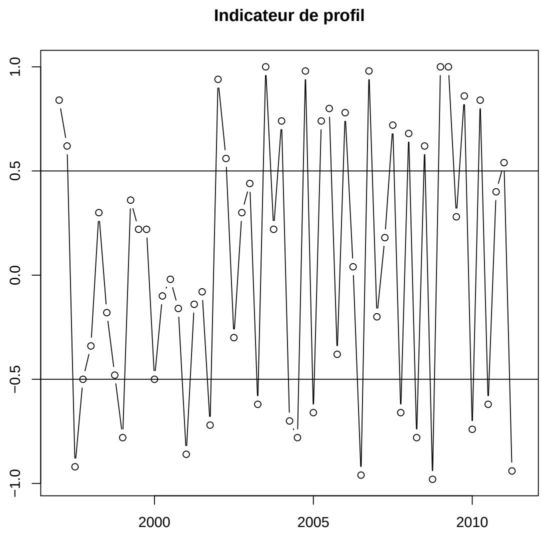
$$\mathbb{P}(\varepsilon_q = 1 | X_q = x_q) = F(\beta x_q)$$

With our previous notations, we can then easily define our directional risk index as :

$$\hat{\mathbb{P}}_q(\hat{\varepsilon}_{q+1} = \varepsilon_{q+1}) = 2(F(\hat{\beta} x_q) - 0,5).$$

We obtain the following index :

Over our historical sample, the directional risk index falled in the "reliable area" (e.g. out of the "area of uncertainty") around 61% of the time. Inside that area, mean of forecast errors $\{\eta_q = 1\}$ also falls below 4% (that is, only 2 misleading forecasts over the period 1997-2010 in this area).



Conclusion

In this article, we studied French GDP directional change forecasts rather than level growth predictions. Quantitative strategies based on business surveys such as Linear Discrimant Analysis can largely surpass a random forecaster performance, with an error rate of 12%. Eventually, using our directional risk index, an economic forecaster can even specify the uncertainty inherent to his directional forecast and conditional on present economic information.

Références

[1] Bardaji, J., Clavel, L., and Tallet, F. Constructing a markov-switching turning point index using mixed frequencies with an application to french business survey data. *OECD Journal : Journal of Business Cycle Measurement and Analysis 5*, 2 (2009), 111–132.

[2] Bardaji, J., Minodier, C., Clavel, L., and Tallet, F. Deux nouveaux indicateurs pour aider au diagnostic conjoncturel en france. *Note de Conjoncture (decembre), INSEE* (2008).

[3] Bessec, M. Etalonnages du taux de croissance du pib français sur la base des enquêtes de conjoncture. *Economie & prévision*, 2 (2010), 77–99.

[4] Biau, G., Bleakley, K., Györfi, L., and Ottucsák, G. Nonparametric sequential prediction of time series. *Journal of Nonparametric Statistics 22*, 3 (2010), 297–317.

[5] Breiman, L. *Classification and regression trees.* Chapman & Hall/CRC, 1984.

[6] Clavel, L., and Minodier, C. A monthly indicator of the french business climate. *Documents de Travail de la DESE-Working Papers of the DESE* (2009).

[7] Cornec, M., and Deperraz, T. Un nouvel indicateur synthétique mensuel résumant le climat des affaires dans les services en france. *Economie et statistique 395*, 1 (2006), 13–38.

[8] Dubois, É., and Michaux, E. Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats. *Economie & prévision*, 1 (2006), 11–28.

[9] Duchêne, S., and Ouvrard, J.-F. La prevision des comptes de la zone euro à partir des enquêtes de conjoncture. *Note de Conjoncture (mars), INSEE* (2011).

[10] Erkel-Rousse, H., and Minodier, C. Do business tendency surveys in industry and services help in forecasting GDP growth ? A real-time analysis on french data. *Documents de Travail de la DESE-Working Papers of the DESE, INSEE* (2009).

[11] Goossens, M., Mittlebach, F., and Samarin, A. *The Latex Companion.* Addison-Wesley, Reading, Massachusetts, 1993.

[12] Hastie, T., Tibshirani, R., and Friedman, J. *The elements of statistical learning : data mining, inference, and prediction.* Springer Verlag, 2009.

[13] Henriksson, R., and Merton, R. On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *The Journal of Business 54*, 4 (1981), 513–533.

[14] Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association 58*, 301 (1963), 13–30.

[15] Leung, M., Daouk, H., and Chen, A. Forecasting stock indices : a comparison of classification and level estimation models. *International Journal of Forecasting 16*, 2 (2000), 173–190.

[16] Maurer, A., and Pontil, M. Empirical bernstein bounds and sample variance penalization. *Arxiv preprint* (2009).

[17] Merton, R. On market timing and investment performance. I. An equilibrium theory of value for market forecasts. *The Journal of Business 54*, 3 (1981), 363–406.

[18] Moore, G. Forecasting short-term economic change, 1983.

[19] national de la statistique et des études économiques (France). Direction des études et synthèses économiques, I., Doz, C., and Lenglart, F. *Factor analysis and unobserved component models : an application to the study of French business surveys.* INSEE, 1996.

[20] Pesaran, M., and Timmermann, A. A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics 10*, 4 (1992), 461–465.

[21] Pesaran, M., and Timmermann, A. How costly is it to ignore breaks when forecasting the direction of a time series ? *International Journal of Forecasting 20*, 3 (2004), 411–425.

[22] Pons, J. The rationality of price forecasts : a directional analysis. *Applied Financial Economics 11*, 3 (2001), 287–290.

[23] Schnader, M., and Stekler, H. Evaluating predictions of change. *The Journal of Business 63*, 1 (1990), 99–107.

[24] Sedillot, F., and Pain, N. Indicator models of real GDP growth in selected OECD countries. *OECD Economics Department Working Papers* (2003).

[25] Sinclair, T., Stekler, H., and Kitzinger, L. Directional forecasts of GDP and inflation : a joint evaluation with an application to Federal Reserve predictions. *Applied Economics 42*, 18 (2010), 2289–2297.

[26] Stekler, H. Are economic forecasts valuable ? *Journal of Forecasting 13*, 6 (1994), 495–505.

[27] Tay, A., and Wallis, K. Density forecasting : a survey.

[28] Wallis, K. Macroeconomic forecasting : a survey. *The Economic Journal 99*, 394 (1989), 28–61.

Appendices

Proof of proposition 1

Notice that :

$$\begin{aligned}\widehat{Cov}(H_t, H_{t-1}) - Cov(H_t, H_{t-1}) &= \Big(\hat{E}(H_t, H_{t-1}) - \hat{E}(H_{t-1})\hat{E}(H_t)\Big) \\ &\quad - (E(H_t, H_{t-1}) - E(H_{t-1})E(H_t)) \\ &= (\hat{E}(H_t, H_{t-1}) - E(H_t, H_{t-1})) \\ &\quad + \hat{E}(H_{t-1})(\hat{E}(H_t) - E(H_t)) \\ &\quad + E(H_t)(\hat{E}(H_{t-1}) - E(H_{t-1}))\end{aligned}$$

Denoting $Z_i := \begin{pmatrix} H_{i-1}H_i \\ H_i \\ H_{i-1} \end{pmatrix}$,we have :

$$\begin{aligned}\widehat{Cov}(H_t, H_{t-1}) - Cov(H_t, H_{t-1}) &= (\hat{E}(Z_t^1) - E(Z_t^1)) + \hat{E}(Z_t^3)(\hat{E}(Z_t^2) - E(Z_t^2)) \\ &\quad + E(Z_t^2)(\hat{E}(Z_t^3) - E(Z_t^3)) \\ &= (1, \hat{E}(Z_t^3), E(Z_t^2))(\hat{E}(Z_t) - E(Z_t))\end{aligned}$$

Assuming that time-dependence is of order 1 at the most for H_t, time-dependency of Z_t is then of order 2 at the most. We can then apply a central limit theorem :

$$\sqrt{Q}\left(\sum_{i=1}^{Q} Z_i - EZ_1\right) \Rightarrow_\infty N(0, \Sigma)$$

with

$$\Sigma = Var(Z_1) + 2\left[cov(Z_1, Z_2) + cov(Z_1, Z_3)\right] = Var(Z_1 + Z_2 + Z_3)$$

and

$$cov(Z_i, Z_j) := sym\left(E\left[Z_i - E(Z_i)\right]\left[Z_j - E(Z_j)\right]'\right)$$

Indeed if we try to find a_T such as $Var(a_T\hat{Z}_T) \rightarrow_{T\infty} \kappa \neq 0$:

$$
\begin{aligned}
Var(a_T\hat{Z}_T) &= \frac{a_T^2}{T^2}Var(\sum_{j=1}^{T} Z_j) \\
&= \frac{a_T^2}{T^2}\mathbb{E}\left(\left[\sum_{j=1}^{T} Z_j - \mathbb{E}\left(\sum_{j=1}^{T} Z_j\right)\right]\left[\sum_{j=1}^{T} Z_j - \mathbb{E}\left(\sum_{j=1}^{T} Z_j\right)\right]^{'}\right) \\
&= \frac{a_T^2}{T^2}\mathbb{E}\left(\left[\sum_{j=1}^{T} (Z_j - \mathbb{E}Z_j)\right]\left[\sum_{j=1}^{T} (Z_j - \mathbb{E}Z_j)\right]^{'}\right) \\
&= \frac{a_T^2}{T^2}\sum_{1\leq i,j\leq T}\mathbb{E}\left((Z_i - \mathbb{E}Z_i)(Z_j - \mathbb{E}Z_j)^{'}\right) \\
&= \frac{a_T^2}{T^2}\left(\sum_{j=1}^{T} Var(Z_i) + \sum_{1\leq i\neq j\leq T}\mathbb{E}\left((Z_i - \mathbb{E}Z_i)(Z_j - \mathbb{E}Z_j)^{'}\right)\right) \\
&= \frac{a_T^2}{T}(Var(Z_1) + \frac{T-1}{T}\left(\mathbb{E}\left((Z_1 - \mathbb{E}Z_1)(Z_2 - \mathbb{E}Z_2)^{'}\right) + \mathbb{E}\left((Z_2 - \mathbb{E}Z_2)(Z_1 - \mathbb{E}Z_1)^{'}\right)\right) \\
&\quad + \frac{T-2}{T}\left(\mathbb{E}\left((Z_1 - \mathbb{E}Z_1)(Z_3 - \mathbb{E}Z_3)^{'}\right) + \mathbb{E}\left((Z_3 - \mathbb{E}Z_3)(Z_1 - \mathbb{E}Z_1)^{'}\right)\right)) \\
&= \frac{a_T^2}{T}\left(Var(Z_1) + \frac{T-1}{T}2cov(Z_1,Z_2) + \frac{T-1}{T}2cov(Z_1,Z_3)\right)
\end{aligned}
$$

Hence we have to choose $a_T = \sqrt{T}$,which leads to :

$Var(\sqrt{T}\hat{Z}_T) \rightarrow_{T\infty} \Sigma := Var(Z_1)+2cov(Z_1,Z_2)+2cov(Z_1,Z_3) = Var(Z_1+Z_2+Z_3)-Var(Z_1+Z_2)$

Denoting $\lambda := (1, E(Z_t^3), E(Z_t^2))^{'}$ central limit theorem gives, under suitable conditions :

$$\sqrt{Q}(\lambda^{'}(\hat{E}(Z_t) - E(Z_t))) \Rightarrow N(0, \lambda^{'}\Sigma\lambda)$$

which is equivalent to :

$$\frac{\sqrt{Q}\lambda^{'}(\hat{E}(Z_t) - E(Z_t))}{\sqrt{\lambda^{'}\Sigma\lambda}} \Rightarrow N(0,1)$$

Finally we get, by replacing $\widehat{\lambda}$ in the following expression :

$$\frac{\sqrt{Q}\hat{\lambda}^{'}(\hat{E}(Z_t) - E(Z_t))}{\sqrt{\hat{\lambda}^{'}\hat{\Sigma}\hat{\lambda}}} \Rightarrow N(0,1)$$

by $\hat{\lambda} := (1, \hat{E}(Z_t^3), \hat{E}(Z_t^2))^{'}$et $\hat{\Sigma} := \widehat{Var}(Z_1) + 2\left[\widehat{cov}(Z_1,Z_2) + \widehat{cov}(Z_1,Z_3)\right]$:

$$\frac{\sqrt{Q}\left(\widehat{Cov}(H_t,H_{t-1}) - Cov(H_t,H_{t-1})\right)}{\sqrt{\hat{\lambda}^{'}\hat{\Sigma}\hat{\lambda}}} \Rightarrow N(0,1)$$

From this previous expression we derive an asymptotic α-level test for the null hypothesis $H_0 := \{Cov(H_t,H_{t-1}) = 0\}$:

$$\left\{\left|\frac{\sqrt{Q}\widehat{Cov}(H_t,H_{t-1})}{\sqrt{\hat{\lambda}^{'}\hat{\Sigma}\hat{\lambda}}}\right| \geqslant q_{1-\alpha/2}\right\}$$

□

Proof of proposition 3

Using Hoeffding's inequality in ([14]) :

$$\begin{aligned}\mathbb{P}(|\frac{1}{Q}\sum_{q}^{Q+h} H_q - \widehat{p_1}| \geq \varepsilon + \tilde{\varepsilon}) &\leq \mathbb{P}(|\frac{1}{Q}\sum_{q} H_q - p_1| \geq \varepsilon) + \mathbb{P}(|p_1 - \widehat{p_1}| \geq \tilde{\varepsilon}) \\ &\leq 2\exp(-2Q\varepsilon^2) + 2\exp(-2N\tilde{\varepsilon}^2)\end{aligned}$$

Let us choose $\varepsilon, \tilde{\varepsilon}$ such as $2\exp(-2Q\varepsilon^2) \leq \alpha/2$ and $2\exp(-2N\tilde{\varepsilon}^2) \leq \alpha/2$, i.e. $\varepsilon = \sqrt{\frac{\ln(2/\alpha)}{2Q}}$et $\tilde{\varepsilon} = \sqrt{\frac{\ln(2/\alpha)}{2N}}$.

□

Proof of proposition 4

For this, recall two useful inequalities :

1. Bennett's inequality ([14]) : with probability $1-\delta$:

$$\left|\bar{H}_n - \mathbb{E}H\right| \leq \sqrt{\frac{2\mathbb{V}(H)\ln(2/\delta)}{n}} + \frac{\ln(2/\delta)}{3n}$$

2. A second inequality in [16] makes the connection with the empirical variance : with probability $1-\delta$, we have :

$$\sigma \leq \hat{\sigma}_n + \sqrt{\frac{2\ln(1/\delta)}{n-1}}$$

Combining these two inequality, we obtain with probability $1-4\delta$:

$$\begin{aligned}|\frac{1}{Q}\sum_{q} H_q - \widehat{p_1}| &\leq |\frac{1}{Q}\sum_{q} H_q - p_1| + |p_1 - \widehat{p_1}| \\ &\leq \sqrt{\frac{2\sigma^2\ln(2/\delta)}{Q}} + \frac{\ln(2/\delta)}{3Q} + \sqrt{\frac{2\sigma^2\ln(2/\delta)}{N}} + \frac{\ln(2/\delta)}{3N} \\ &\leq \sqrt{\frac{2\ln(2/\delta)}{Q}}\left(\hat{\sigma}_N + \sqrt{\frac{2\ln(1/\delta)}{N-1}}\right) + \frac{\ln(2/\delta)}{3Q} \\ &\quad + \sqrt{\frac{2\ln(2/\delta)}{N}}\left(\hat{\sigma}_N + \sqrt{\frac{2\ln(1/\delta)}{N-1}}\right) + \frac{\ln(2/\delta)}{3N} \\ &\leq \hat{\sigma}_N\sqrt{2\ln(2/\delta)}\left(\frac{1}{\sqrt{Q}} + \frac{1}{\sqrt{N}}\right) + \frac{7\ln(2/\delta)}{3(N-1)} + \frac{\ln(2/\delta)}{3Q} + \frac{2\ln(2/\delta)}{\sqrt{Q(N-1)}}\end{aligned}$$

□