RISK

RETURN

RESEARCH

Kernel Regression

The goal of regression is to analyze the relationship between a response variable and explanatory variable(s). Ordinary least-squares regression is the most common and most computationally efficient regression technique, but its underlying hypotheses are often too restrictive (especially in finance).

There are dozens of other regression techniques, each one better suited for certain situations. In our field, we often want to regress forward returns of an asset (Y) against an indicator (X), which we believe could influence the asset's returns.

In this context, it is often the case that forward returns are often extremely noisy (especially when we use short-term forward returns). Additionally, the relationship between the indicator and asset returns is unlikely to be linear (or even parametric) in most cases.

Kernel regressions are especially well suited for noisy data. Additionally, they do not impose a linear (or parametric) relationship between variables.

Cyril Castelli, CEO
cyril.castelli@rcube.com

Paul Buigues
paul.buigues@rcube.com

Stéphane Alloiteau
stephane.alloiteau@rcube.com

Remi Takase
remi.takase@rcube.com

phone
+(33) 1 55 04 04 42

email
info@rcube.com

RCUBE WHITE PAPER 17.10.2012

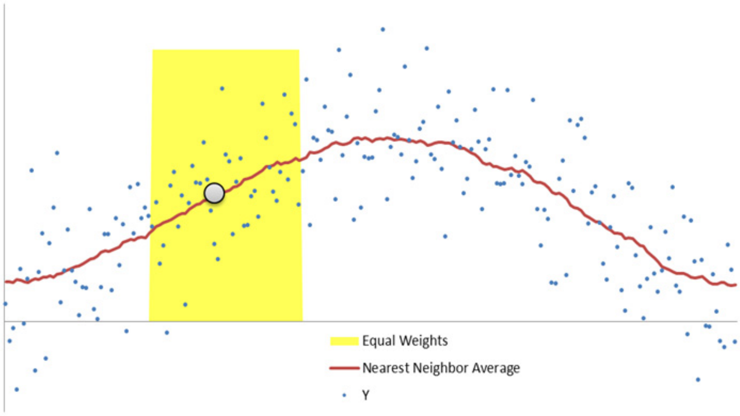
Contents

www.rcube.com
39 rue Cambon

1) What is kernel regression?

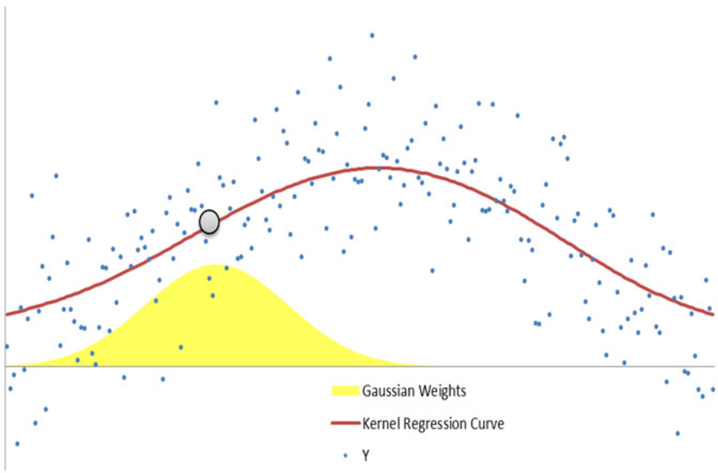
We want to find a smooth, continuous function that fits the expected forward returns while avoiding overfitting.

One way to find this function is to use for each observation the average value of its nearest neighbors. However, by estimating the conditional expectation this way, the resulting estimated function will always be discontinuous:



The discontinuities are due to the equal-weighting of observations, and to the fact that observations are treated in a binary way (they are either inside or outside the neighborhood).

In order to smooth the curve, one solution consists in assigning weights to each observation, depending on its distance from the target point. This type of regression is called a kernel regression (or Nadaraya-Watson regression), the "kernel" being the function used to assign weights to each observation. The most widely used kernel is the Gaussian kernel[1].



---

[1] The formula for the regression curve is given by: $\hat{f}(x) = \frac{\sum_{i=1}^{n} Y_i K(\frac{x-X_i}{\lambda})}{\sum_{i=1}^{n} K(\frac{x-X_i}{\lambda})}$

where $K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$ for a Gaussian kernel. More generally, kernels are non-negative functions, which are symmetric and have an integral equal to 1.

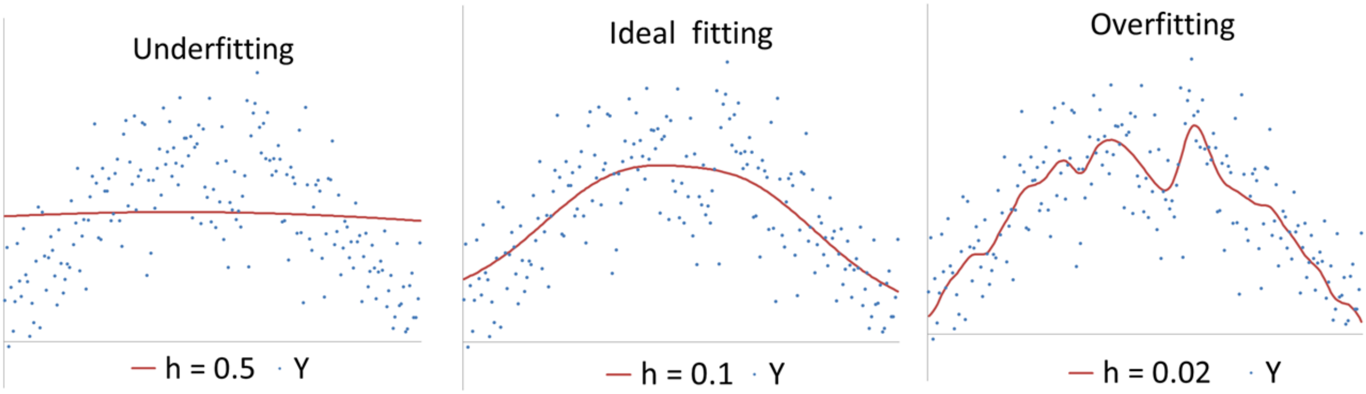
The degree of smoothness: the bandwidth λ

Kernel regression belongs to "non-parametric" statistics, in the sense that it does not impose a predetermined form to the estimated function.

Nevertheless, there is one parameter: the bandwidth (λ). When a Gaussian kernel is used, the bandwidth represents its standard deviation. It defines the width of the Gaussian curve, and controls how quickly the weights decrease from the target point.

The estimator's smoothness changes according to the bandwidth: when λ is very small, the resulting estimator will be very accurate, but it will be too close to the data (overfitting). On the other hand, if the value of the bandwidth is too large, we will obtain an over-smoothed regression curve (underfitting). In our case, we will often take $\lambda = h * number\ of\ observations$, with $h = 0.1$ or $h = 0.2$.

The following examples illustrate the importance of choosing a good bandwidth:



There are methods that automatically find a trade-off between the variance and the bias, but they are computational greedy and the results are not necessarily relevant for us.

Note: In practice, we will generally use the rank of an indicator rather than the indicator itself for X. This is due to the fact that most of the time, the indicator's values are not equally distributed, which can lead to sparsely populated areas in the indicator's axis, and therefore to a less robust estimator.

2) Kernel regression example: VXX vs. VIX term structure

One of the strongest relationships we have identified for volatility trading is the impact of the volatility term structure on the returns of a long volatility strategy.

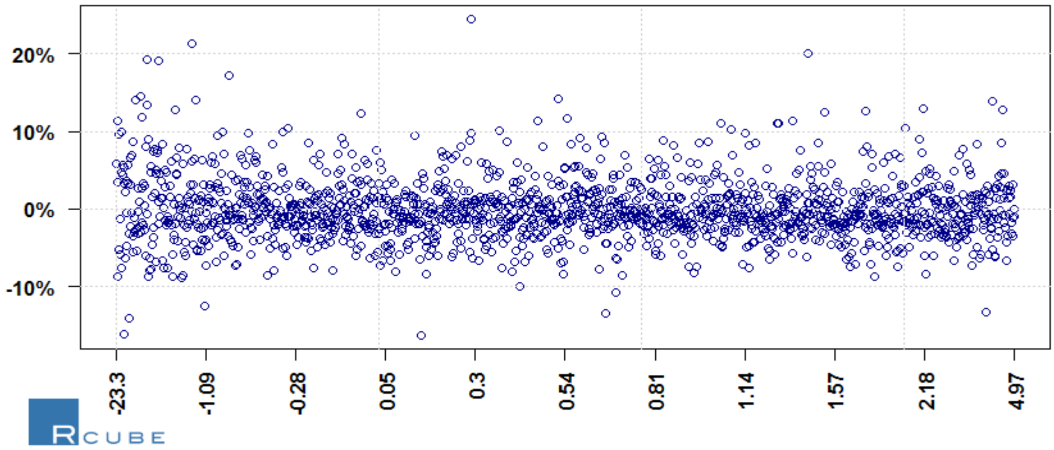
In the case of the VIX, we would therefore want to perform a regression of Y against X, with:

X = VIX Term Structure, i.e. the front-term future minus the spot Value (UX1 – VIX)

Y = VIX Short-Term Futures Index (SPVXSTR) returns, traded since 2009 as an ETF as “VXX”[2]

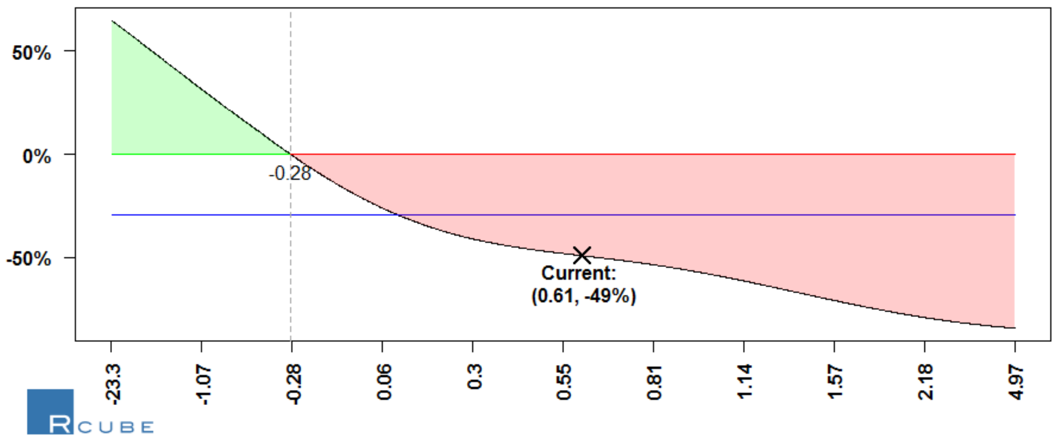
The following graph plots daily forward returns of the VXX and the VIX term structure (UX1 – VIX). Any potential relationship is unlikely to be discernible to the eye:

DAILY FORWARD RETURNS OF VXX NAV VS. VIX TERM STRUCTURE (UX1-VIX)



However, a simple kernel regression analysis produces a clear pattern:

KERNEL REGRESSION ANALYSIS
FORWARD ANNUALIZED RETURNS OF VXX VS. VIX TERM STRUCTURE (UX1-VIX)

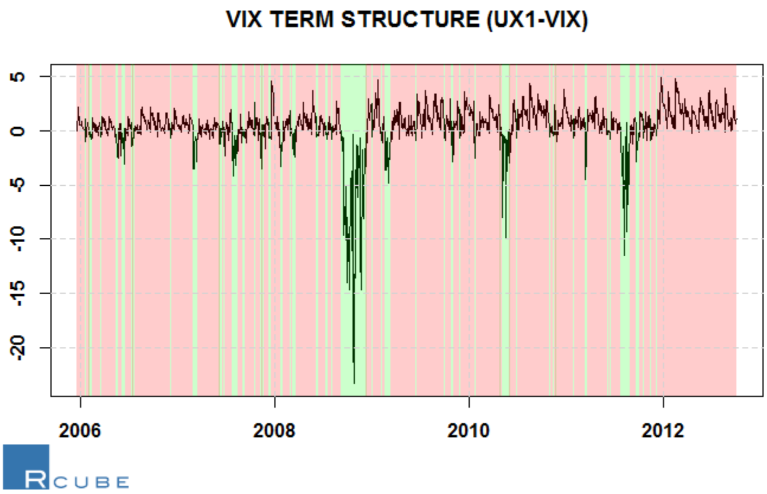


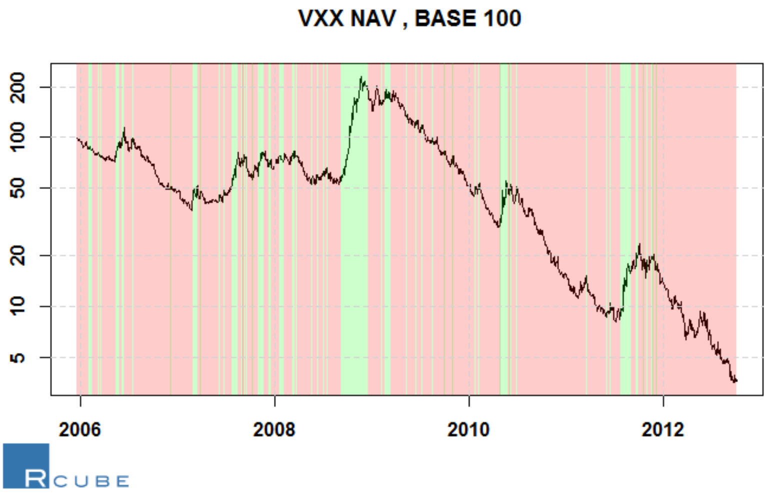
[2] The VIX Short-Term Futures Index consists in being long in a combination of 1- and 2- month futures contracts, and selling every day a fraction of the 1-month to buy a fraction of the 2-month contract.

The regression curve represents the expected returns of the asset (the VXX ETF) for a given value of the indicator (UX1 - VIX). The picture also indicates the current value of the indicator (as of October 12th, 2012) and its expected forward returns. The regression curve highlights in different colors areas of the explanatory variable where returns are positive (green) or negative (red) *on average*. The blue line corresponds to the average of VXX returns.

Backtesting a strategy based on kernel regression analysis

The intersection that we found between the regression line and the X axis (i.e. the expected forward return = 0%) can be used as entry/exit thresholds in a backtest[3]. We therefore test a strategy that goes long on VXX when the VIX term structure is lower than -0.28%, and short otherwise.

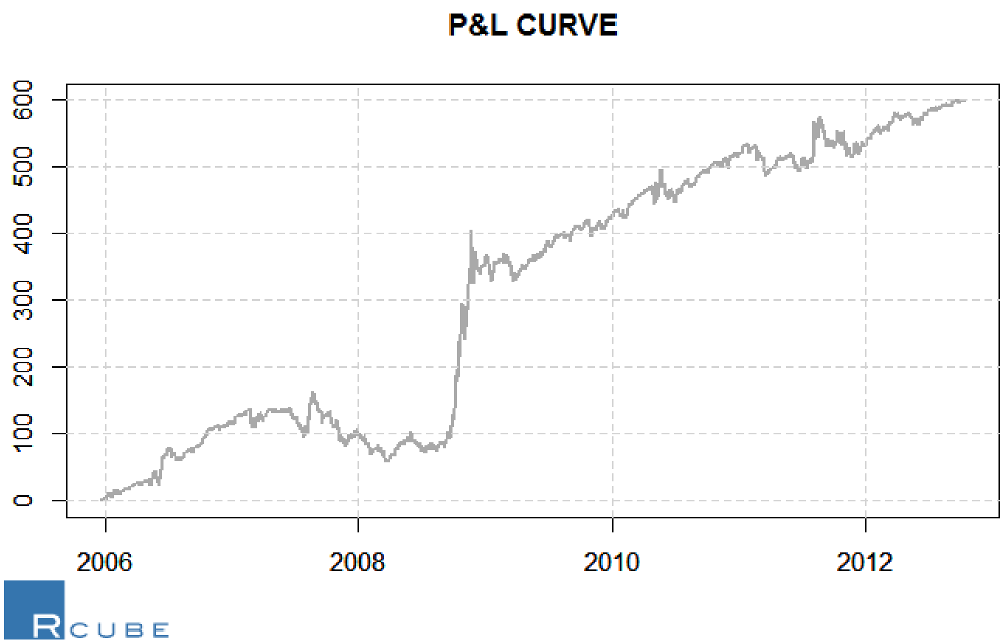




[3] We are fully aware of the fact that this analysis is performed in-sample, and that there is no reason for the threshold to remain at the exact same location until the end of times. Additionally, it is worth noting that a threshold is not optimal from a total return point of view (to get an optimal – and likely overfitted - location, we would have to use an isotonic regression – see our Rcube White Paper – “Cumulative Curves, Fractiles and Monotonic Regressions)

On a trade-by-trade basis (where we invest a constant amount of 100 at each trade), the associated strategy generated an average performance of 85.34% and a Sharpe Ratio of 1.17:

	Trades		Statistics
Total trades	211	**Cagr**	85.34%
Total win trades	88	**Vol**	73.18%
Hit miss ratio	41.71%	**Sharpe**	1.17
Avg performance	2.75%		
Win performance	12.51%		
Loss performance	-4.22%		
Win loss ratio	2.96		
Avg trade duration	12 days		
Avg win trade duration	21 days		
Avg loss trade duration	5 days		



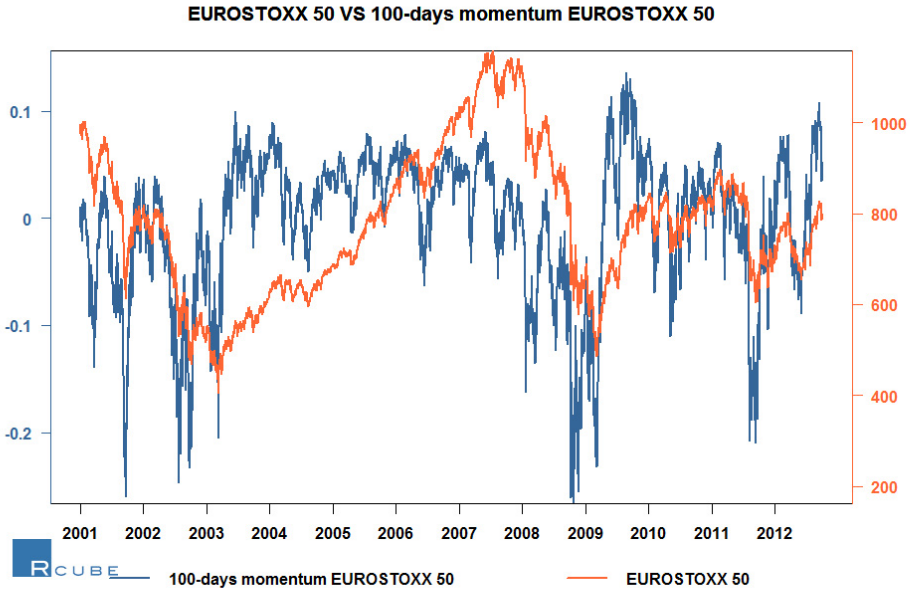
While it is unlikely for the strategy to generate the same extraordinary performances going forward, we can conclude that it is worth looking at the volatility term structure before entering a directional volatility position.

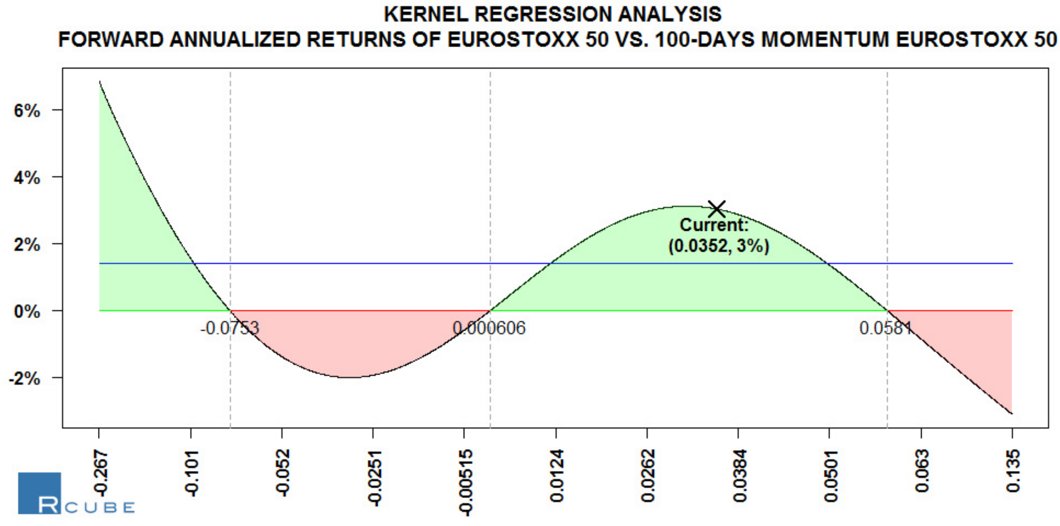
3) Kernel regression example: Eurostoxx 50 momentum

In addition to dealing with noisy data, kernel regressions also facilitate the discovery of non-monotonic relationships between two variables. Stock index momentums provide a good illustration of this type of relationship. We will therefore perform a regression of Y against X, with:

X = Eurostoxx 50 momentum, defined as the ratio minus 1 between the Eurostoxx' spot value and its 100-days EMA

Y = Eurostoxx 50 returns





The curve shows an interesting property of the relationship which would not have been observable with an OLS/isotonic regression. The asset is clearly mean-reverting for extreme values (both boundaries) of the momentum and otherwise trend-following.

This seems to validate the saying that "fast moves mean-revert, while slow moves persist". Both mean-reversion and trend-following coexist in the market, depending on the speed of the moves.

4) Conclusion

Kernel regression analysis can be considered as an enhancement over bar plots. In bar plots, we have to discretize an explanatory variable to calculate expected returns over each interval. In kernel regression, we obtain a smooth and continuous function, which provides an expected return for each value of the explanatory variable - although we should always refrain from placing too much confidence in expected forward returns.

Another interesting advantage of kernel regression is its ability to detect non-linear but relevant relationships in the data.

We will therefore use kernel regression quite often in our upcoming research notes.

Research disclaimer

****The portfolios mentioned within this research material are hypothetical, no actual investments have occurred****

The model portfolio is a hypothetical portfolio and is not intended to reflect actual trading performance. As a hypothetical model, the portfolio does not necessarily represent a portfolio of securities that the Firm holds or trades. Any references to losses, gains or position with respect to the portfolio represent hypothetical not actual losses, gains or positions.
Any financial promotion contained herein has been approved and issued by Rcube SAS;
This material has been prepared for distribution to market professional and institutional investor clients. Recipients who are market professionals or institutional investor clients of Rcube SAS should seek independent financial advice prior to making any investment decision based on this material. This material does not provide individually tailored investment advice. No part of this material may be copied, photocopied, or duplicated in any form, by any means, or redistributed without Rcube SAS prior written consent. Further information on any of the securities mentioned herein may be obtained upon request. Any opinions expressed herein are those of the individual author and may differ from others opinions expressed. This communication is being provided for informational purposes only and not intended as a recommendation or an offer or solicitation for the purchase or sale of any security referenced herein. Rcube SAS has obtained all market prices, data and other information from sources believed to be reliable although its accuracy or completeness cannot be guaranteed. Such information is subject to change without notice. If this publication has been distributed by electronic transmission, such as e-mail, then such transmission cannot be guaranteed to be secure or error-free as information could be intercepted, corrupted, lost, destroyed, arrive late or incomplete, or contain viruses. The sender therefore does not accept liability for any errors or omissions in the contents of this publication, which may arise as a result of electronic transmission. If verification is required, please request a hard-copy version. Rcube SAS is registered under 529 778 177 in the PARIS'R.C.S