ESTIMATION AND TESTING FOR TWO-DIMENSIONAL DIFFUSIONS IN FINANCE: EXPLORING A SEMIPARAMETRIC PROPOSAL

DISSERTAÇÃO SUBMETIDA À CONGREGAÇÃO DA ESCOLA DE PÓS-GRADUAÇÃO EM ECONOMIA (EPGE) PARA OBTENÇÃO DO GRAU DE MESTRE EM ECONOMIA POR CRISTIAN HUSE

RIO DE JANEIRO, RJ
NOVEMBRO, 1999
Für meine Großeltern,

Käte und Rudolf, Erica und Gerhard
ACKNOWLEDGEMENTS

I have no words to thank Renato Flôres for accepting me as a pupil in a moment of difficulty and disbelief concerning academic life. With his help and influence, I have chosen my professional option, besides getting conscious not of the nonexistence, but of the negligibility, of certain aspects of the career. His courses, his orientation, our conversations and his very particular way of viewing the field of study were fundamental to my background. Needless to say, the remaining errors that might appear in this work should not be blamed on him, being solely mine.

I would also like to thank those teachers whose influence, if not explicitly mentioned, is clearly reflected in the way I face life and, of course, in the following pages. I hope this influence remains in the many I intend to write.

I would also like to thank Irina Potapenko, Mihail Lermontov and Luis Laurencel for computational support and exchange of ideas during this work.

Financial support from CNPq/Brasil is also gratefully acknowledged.

Finally, I cannot forget to thank my family and friends for all support. In fact, it is so present and intense that it is almost not perceived.
1. Introduction

Derivatives are assets whose value depends on other variables. These underlying variables can also be assets, though not necessarily, as in the case of interest rates. Futures, forwards, swaps, caps, floors, options on stocks, or on bills, indexes, currencies, futures and interest rates are examples of derivatives. They can be traded both in exchanges or over the counter, and their volume is, in general, much larger than that of the underlying assets.

Fundamental landmarks in the pricing of derivatives are the Black and Scholes (1973) and Merton (1973) papers, developed in continuous-time and using arbitrage arguments, which pioneered a path followed by Vasicek (1977), Cox, Ingersoll and Ross (1985a, b) and Hull and White (1987), among others. Though the pricing framework in continuous time is generally much more tractable and elegant than the alternative discrete approximations, the empirical literature did not follow its theoretical counterpart. Typically, the estimation of derivative pricing models abandons the continuous time environment, restricting itself to the data available, which are in discrete time.

Let \( \{x_t, t \geq 0\} \), the underlying process of the price of a given asset, be a diffusion represented by the Itô stochastic differential equation

\[
\frac{dx_t}{x_t} = \mu(x_t) \, dt + \sigma(x_t) \, dW_t
\]

where \( \{W_t, t \geq 0\} \) is a standard Brownian motion. The functions \( \mu(.) \) and \( \sigma(.) \) are, respectively, the drift (instantaneous mean) and the volatility (instantaneous variance) of the process.

Various specifications for the drift and volatility of diffusions in finance have been proposed. Vasicek (1977) models interest rates, \( r \), using an Ornstein-Uhlenbeck process,
where \((\mu(r), \sigma(r)) = (\beta(\alpha-r), \sigma)\), a drift with the mean-reversion property and a constant instantaneous variance. Cox, Ingersoll and Ross (1985) develop a general equilibrium model resulting in a process where \((\mu(r), \sigma(r)) = (\beta(\alpha-r), \sigma \sqrt{r})\), a drift with the mean-reversion property and an instantaneous variance as a linear function of the short term interest rate.

With discrete data, Lo (1988), in a pioneering paper, proposed an estimation approach based on the method of maximum-likelihood, whose drawback is, except for very particular cases, to require the numerical solution of a partial differential equation for each maximum-likelihood iteration. Nelson (1990) analysed the behavior of discrete approximations when the interval between the observations goes to zero. Duffie and Singleton (1993) and Gourieroux, Monfort and Renault (1993) proposed the estimation of diffusions by simulation – given parameter values, sample paths are simulated, and their moments should be as close as possible to the sample moments. Finally, a commonly used method consists in parametrizing the functions \(\mu(.)\) and \(\sigma^2(.)\) and then discretize the model before estimating it.

The approach proposed by Aït-Sahalia (1996) seeks to reconcile both the theoretical and empirical literature in option pricing. Although the data are discrete, it does not resort to discretizations of the model. Firstly, one parametrizes the drift, for instance, which guarantees the identification of the model and makes it possible not to restrict the volatility in any way. After imposing the parametrization, one proceeds to the following steps:

1. Estimate the drift of the diffusion, using Ordinary Least Squares (OLS);
2. Estimate nonparametrically the marginal density of the process, using kernel smoothers;
3. Given the estimated values of the drift and the marginal density function of the process, obtain the nonparametric estimator of the volatility as a solution to the Kolmogorov forward equation;
4. Using the volatility estimates as input, reestimate (i) the drift parameters using Feasible Generalized Least Squares (FGLS), (ii) the volatility of the process.
Aït-Sahalia enumerates three reasons which make his approach attractive:

(i)  The importance of the volatility in derivatives pricing;

(ii) The difficulty of forming an a priori idea of the functional form of the instantaneous volatility, as it is not observed;

(iii) The availability of long time-series of daily data of spot interest rates.

The purpose of this thesis is to explore the possibilities of bivariate extensions of Aït-Sahalia (1996)'s framework. Bivariate diffusions appear in two-factor models and are many times treated independently. It would be interesting to have a powerful estimating method which would allow for investigating, and testing, different relationships among the two univariate processes.

The essay is organized as follows. Section 2 briefly recovers Aït-Sahalia's univariate approach, while the following section analyses the bivariate case. Section 4 proposes several semiparametric estimators for the covariance between the processes. The next section applies the bivariate approach to a pair of assets: the main stock indexes of Brazil and Argentina. Section 6 concludes.

The main message of this thesis is that extending Aït-Sahalia (1996)'s framework to a multivariate setting is by no means straightforward. First, the Itô's and Fokker-Planck's frameworks do not coincide in higher dimensions, as opposed to the univariate case. As a matter of fact, the Fokker-Planck equation turns out to be more adequate in the bivariate case than the Itô representation. Second, even when imposing sensible assumptions on the drift and volatility functions, one cannot obtain a direct generalization of the univariate method. In spite of these issues, the method might be interesting as an exploratory technique for uncovering certain relationships between the two processes, without imposing a fully parametric structure. Further work is however needed for both rigorously establishing the
asymptotic properties of the various possible estimators, as well as to acquire a better grasp of relevant potential applications.

2. Univariate Semiparametric Estimation of Diffusions

Ait-Sahalia (1996) considers the univariate Kolmogorov Forward Equation (Karlin and Taylor (1981), p. 219), or the Fokker-Planck (FP) equation, which describes the transition densities of continuous-time Markov processes without jump:

\[
\frac{\partial}{\partial t} f(x,t;y,t') = -\frac{\partial}{\partial x} (\mu(x)f(x,t;y,t')) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(x)f(x,t;y,t'))
\]

where:
- \( f(x,t;y,t') \) := transition density from point \( (y,t') \) to \( (x,t) \);
- \( \mu(x) \) := drift of the process;
- \( \sigma^2(x) \) := volatility of the process.

A first assumption concerns the drift, parametrized as in Vasicek (1977), inspired in the Ornstein-Uhlenbeck process, which has the mean-reverting property. Parametrization of the drift is fundamental to the identification of the pair \( (\mu,\sigma^2) \): imposing no restriction on the pair makes it impossible to distinguish it from the pair \( (a\mu,a\sigma^2) \), where \( a \) is a constant, when considering a discrete sample with fixed time-intervals. Supposing a general parametrization \( (\mu(x,\theta)) \), one obtains:
The second assumption is the stationarity of the process – or rather, that it has converged to a steady state i.e. to a stationary distribution - which implies \( \frac{\partial}{\partial t} f(x,t;y,t') = 0 \) and \( \pi(x,t) = \pi(x) \), where \( \pi(\cdot) \) is the marginal density of the process. Multiplying both sides of (2.2) by \( \pi(y) \) and integrating with respect to \( y \), one obtains,

\[
\frac{\partial}{\partial x} \left( \mu(x,\theta) \right) \int f(x,t;y,t') \pi(y) dy = \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( \sigma^2(x) \int f(x,t;y,t') \pi(y) dy \right)
\]

Using again the assumption of stationarity, it is possible to rewrite (2.3) as

\[
\frac{d}{dx} (\mu(x,\theta) \pi(x)) = \frac{1}{2} \frac{d^2}{dx^2} (\sigma^2(x) \pi(x)).
\]

Integrating (2.4) one obtains

\[
\mu(x,\theta) \pi(x) = \frac{1}{2} \frac{d}{dx} \left( \sigma^2(x) \pi(x) \right),
\]

and integrating again with respect to \( x \) and using the boundary condition \( \pi(0) = 0 \)

\[
\sigma^2(x) = \frac{2}{\pi(x)} \int_0^x \mu(u,\theta) \pi(u) du.
\]
It is then possible to write the volatility as an explicit function $\phi(\theta ; \pi(x))$ of the marginal density and the parameter vector characterizing the drift. If these two objects are estimated, a semiparametric estimate of the volatility function can be obtained as:

\begin{equation}
\hat{\sigma}^2(x) = \phi(\hat{\theta} ; \hat{\pi}(x))
\end{equation}

Aït-Sahalia (1996) shows that the diffusion function can be identified from the marginal density of the process at stake and from the vector parameter estimated by OLS. Moreover, the estimator is shown to be pointwise consistent and asymptotically normal.

3. **A Bivariate Generalization**

3.1 **The Bivariate Fokker-Planck Equation**

Consider now the bivariate Fokker-Planck equation, already with a parametrization on the drift:

\begin{equation}
\frac{\partial}{\partial t} f(x,t; y,t') = -\sum_{i=1}^{2} \frac{\partial}{\partial x_i} (\mu_i(x,\theta) f(x,t; y,t')) + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial^2}{\partial x_i \partial x_j} (b_{ij}(x) f(x,t; y,t'))
\end{equation}

where

$f(x,t; y,t') := \text{transition density from point (y,t') to (x,t)}$;

$\mu_i(x,\theta), i=1,2 := \text{drifts of the two processes;}$
First, it is worth to stress the correspondence, at least locally, between the Itô representation and the Fokker-Planck equation (Gardiner (1990), chapter 3). The bivariate version of the Itô stochastic differential equation is:

\[
\begin{bmatrix}
    dx_{1t} \\
    dx_{2t}
\end{bmatrix} = \begin{bmatrix}
    \mu_1(x_1, x_2) \\
    \mu_2(x_1, x_2)
\end{bmatrix} dt + \begin{bmatrix}
    \sigma_{11}(x_1, x_2) & \sigma_{12}(x_1, x_2) \\
    \sigma_{21}(x_1, x_2) & \sigma_{22}(x_1, x_2)
\end{bmatrix} \begin{bmatrix}
    dW_{1t} \\
    dW_{2t}
\end{bmatrix}
\]

where \( x = (x_1, x_2) \), and \( \{(W_{it})_{i=1,2}, t \geq 0\} \) is a standard two-dimensional Brownian motion. The functions \( \mu_i(\cdot) \) and \( \sigma_i^2(\cdot), i = 1,2 \) are, respectively, the drift and the "volatility" of each process, and \( \sigma_{ij}(\cdot), i, j = 1,2, i \neq j \) are instantaneous "covariances" between the processes.

The relation between the volatilities in both representations is \( B = \Sigma \Sigma' \), where \( B \) is the FP volatility matrix, and \( \Sigma \) is the Itô volatility matrix. Thus the relation between (3.1) and (3.2), i.e. the two-dimensional case, is

\[
\begin{align*}
    b_{11}(x) &= \sigma_{11}^2(x) + \sigma_{12}^2(x) \\
    b_{12}(x) &= \sigma_{21}(x) = \sigma_{11}(x)\sigma_{21}(x) + \sigma_{12}(x)\sigma_{22}(x) \\
    b_{22}(x) &= \sigma_{21}^2(x) + \sigma_{22}^2(x)
\end{align*}
\]

To obtain an analytical solution for the bivariate version of the FP equation in the spirit of Aït-Sahalia (1996), one needs some assumptions that make the analysis more than a simple extension of the earlier method. First of all, we assume stationarity of the process, and notice that \( b_{12} = b_{21} \), which means that the "covariances" of the FP representations are equal.
We then introduce the following additional hypothesis: each drift depends only on its underlying process, which means,

\begin{equation}
\mu_i(x, \theta) = \mu_i(x_i, \theta), \ i = 1, 2
\end{equation}

Multiplying both sides of (3.1) by the joint density \( \pi(y_1, y_2) = \pi(y) \), recalling that the stationarity assumption makes the left side of (3.1) equal to zero and integrating with respect to \( y \) one obtains:

\begin{equation}
\sum_{i=1}^{2} \left( \frac{\partial}{\partial x_i} (\mu_i(x_i; \theta) \pi(x)) \right) = \sum_{i=1}^{2} \left( \frac{\partial^2}{\partial x_i^2} (b_{ii}(x) \pi(x)) + \frac{\partial^2}{\partial x_i \partial x_j} (b_{ij}(x) \pi(x)) \right).
\end{equation}

Integrating now with respect to \( x_1 \) and \( x_2 \):

\begin{equation}
\sum_{i,j=1; i \neq j}^{2} \mu_i(x_i; \theta) \int_{x_j} \pi(x) dx_j = \frac{1}{2} \sum_{i,k=1; i \neq k}^{2} \frac{\partial}{\partial x_i} \int_{x_k} b_{ii}(x) \pi(x) dx_k + b_{ij}(x) \pi(x).
\end{equation}

where all integration constants were set to zero. In fixed-income analysis, it is quite natural to assume \( \pi_1(0) = \pi_2(0) = 0 \). Intuitively, this means assigning a probability zero to the event \{nominal interest rate = 0\}. When considering stock returns, for example, the integration interval considered is \( [x_1^{\text{min}}, x_2^{\text{max}}] \) and an analogous hypothesis is \( \pi_1(x_1^{\text{min}}) = \pi_2(x_2^{\text{min}}) = 0 \). Then one obtains integration constants equal to zero once again.

If, in addition, each “variance” of the FP representation depends only on its own process,
then it is possible to write the FP "covariance" explicitly:

\[
(3.8) \quad b_{12}(x) = \frac{1}{\pi(x)} \left[ \sum_{i=1}^{3} \mu_i(x_i, \theta) \left[ \int_{x_i} \pi(x) dx_j - \frac{1}{2} \sum_{i=1}^{3} \frac{\partial}{\partial x_i} b_{ii}(x_i) \pi_i(x_i) \right] \right]
\]

However, it is still necessary to identify the system (3.3), relating the FP and the Itô volatilities.

3.2 Parametrizing Itô Volatilities

3.2.1 A General Setting

The correspondence between the FP and Itô representations, concerning their volatilities, is much simpler in the univariate case than in the bivariate one. While the FP equation is more convenient in operational terms when developing our estimation procedure, the Itô representation has an intuitive appeal, especially when considering a continuous-time counterpart of a covariance matrix. The aim of this section is to analyse (3.6) and (3.8), according to various assumptions imposed on the perhaps most natural Itô representation (3.2).

The most commonly used univariate volatility parametrizations in finance nowadays are those of Vasicek (1977) and Cox-Ingersoll-Ross (1985), CIR, given by, respectively, \( \sigma^2(x) = \sigma^2 \) and \( \sigma^2(x) = \sigma^2 x \). These volatilities could be incorporated into the model when passing
from the FP to the Itô representation. We must however, as said before, identify the system (3.3). One possibility is to assume that $\sigma_{21} = 0$, what gives:

$$
\begin{align*}
  b_{11}(x) &= \sigma_{11}^2(x) + \sigma_{12}^2(x) \\
  b_{12}(x) &= b_{21}(x) = \sigma_{12}(x)\sigma_{22}(x) \\
  b_{22}(x) &= \sigma_{22}^2(x)
\end{align*}
$$

(3.9)

If, for instance, (3.7) is also imposed, this would additionally imply that

$$
\sigma_{22}(x) = \sigma_{22}(x_2)
$$

(3.10) $\sigma_{11}^2(x) + \sigma_{12}^2(x)$ is independent of $x_2$.

One simple way of fulfilling this second condition (i.e. (3.10)) is to make:

$$
\sigma_{11}(x) = \sigma_{11}(x_i) \quad \text{and} \quad \sigma_{12}(x) = \sigma_{12}(x_i).
$$

(3.11)

As it will be shown below, (3.10) and (3.11) are somewhat stringent conditions and make the procedure more useful for a testing rather than an estimation purpose.

Another idea would be to solve system (3.9) for the $\sigma$'s, obtaining:

$$
\begin{align*}
  \sigma_{11}^2(x) &= b_{11}(x) - \frac{b_{12}^2(x)}{b_{22}(x)} = \frac{b_{11}(x) b_{22}(x) - b_{12}^2(x)}{b_{22}(x)} \\
  \sigma_{12}(x) &= \frac{b_{12}(x)}{\sqrt{b_{22}(x)}} \\
  \sigma_{22}^2(x) &= b_{22}(x)
\end{align*}
$$

(3.12)
Now, again, (3.7) may be imposed but the above system shows clearly that, in principle, both $\sigma_{11}(x)$ and $\sigma_{12}(x)$ will depend on the two components of vector $x$.

All the above assumptions are not, however, sufficient to obtain an analytical solution for either (3.6) or (3.8). As a consequence, compared to the univariate case, one needs additional parametric assumptions. One interesting parametrization, which has also an intuitive appeal, consists in imposing functional forms on the variances of the Itô representation, i.e., on the diagonal terms of the instantaneous covariance matrix of the Itô representation. In particular, consider those variances taking the functional forms of the volatilities of the Vasicek and CIR models; this will make it possible to write explicitly the covariance of the Itô representation as a function of the drifts and variances of both processes, of the joint density of the process and of the marginal density of each process. We shall now explore many forms of this specification.

### 3.2.2 The Double Vasicek Model

Consider the volatilities of the Itô representation. Assume, besides the identification condition $\sigma_{21} = 0$, that they are parametrized as constants, such as in the Vasicek model

\[(3.13) \quad \sigma_{11}(x_i) = C_1 \]
\[\sigma_{22}(x_i) = C_2 \]

These assumptions, together with (3.10), allow us to rewrite (3.9) as

\[b_{11}(x_i) = \sigma_{11}^2(x_i) + \sigma_{12}^2(x_i) = C_1^2 + \sigma_{12}^2(x_i)\]
Inserting (3.14) in (3.8), one obtains a nonlinear differential equation (NDE) with variable coefficients

\begin{equation}
A_1\sigma_{12}(x_1) + A_2\sigma_{13}(x_1)\sigma_{12}'(x_1) + A_3\sigma_{12}^2(x_1) + A_4 = 0
\end{equation}

where

\begin{align}
A_1 &= C_2\pi(x_1, x_2) \\
A_2 &= \pi_1(x_1) \\
A_3 &= \frac{1}{2}\frac{d\pi_1(x_1)}{dx_1} \\
A_4 &= \frac{1}{2}\sum_{i=1}^{2} C_i^2 \frac{d\pi_i(x_1)}{dx_i} - \sum_{i=1}^{2} \mu_i(x_1, \theta) \int_{x_i} \pi(x)dx_j
\end{align}

In spite of the fact that the equation above shows that, once obtained the vector parameter \( \theta \), and constants \( C_1 \) and \( C_2 \) it is possible to identify the covariance between the processes from the joint density \( \pi(., .) \) and the marginal densities \( \pi_1(.) \) and \( \pi_2(.) \), by hypothesis, the solution to (3.15) should be a function of \( x_1 \) only. Nevertheless, inspection of (3.16) and (3.19) shows that these elements are functions of the whole vector \( x \), nothing a priori guaranteeing that the solution, in a given case, will be independent of the \( x_2 \) values. This fact makes this specification, within the context of our proposal, more suitable for a testing procedure rather than for estimation purposes.
3.2.3 The Double Cox-Ingersoll-Ross Model

In spite of its popularity, the Vasicek model has some undesirable features, like the occurrence of processes with negative interest rates. The CIR model overcomes this problem by the convenient specification of the volatility function.

Consider once again equations (3.9). Assume, besides the identification condition $\sigma_{21} = 0$, that the volatilities are parametrized such as in the CIR model:

\begin{align}
\sigma_{11}(x_1) &= C_1 \sqrt{x_1} \\
\sigma_{22}(x_2) &= C_2 \sqrt{x_2}
\end{align}

These assumptions, together with (3.10), allow us to rewrite (3.9) as

\begin{align}
b_{11}(x_1) &= \sigma_{11}^2(x_1) + \sigma_{12}^2(x_1) = C_1^2 x_1 + \sigma_{12}^2(x_1) \\
b_{12}(x) &= \sigma_{12}(x_1)\sigma_{22}(x_2) = C_2 \sqrt{x_2} \sigma_{12}(x_1) \\
b_{22}(x_2) &= \sigma_{22}^2(x_2) = C_2^2 x_2
\end{align}

Inserting (3.21) into (3.8), one obtains a NDE with variable coefficients formally similar to (3.15):

\begin{align}
B_1 \sigma_{12}(x_1) + B_2 \sigma_{12}(x_1) \sigma_{12}'(x_1) + B_3 \sigma_{12}^2(x_1) + B_4 = 0
\end{align}

where
(3.23) \[ B_1 = C_2 \sqrt{x_2} \pi(x_1, x_2) \]

(3.24) \[ B_2 = \pi_i(x_i) \]

(3.25) \[ B_3 = \frac{1}{2} \frac{d\pi_i(x_i)}{dx_i} \]

(3.26) \[ B_4 = \frac{1}{2} \sum_{i=1}^{2} C_i^2 x_i \frac{d\pi_i(x_i)}{dx_i} + \frac{1}{2} \sum_{i=1}^{2} C_i^2 \pi_i(x_i) - \sum_{i=1}^{2} \mu_i(x_i, \theta) \int \pi(x) dx_j \]

The equations above bear the same attributes and the same problem of those from the previous specification, so that the same comment applies.

### 3.3 Parametrizing FP Volatilities

We explore now the combination of (3.12) with (3.7). The diffusion coefficients \( b_{ii}(\cdot) \), \( i=1,2 \), could, for instance, be specified in a “CIR fashion” as:

(3.27) \[ b_{ii}(y_i) = k_i y_i, \quad i = 1, 2 \]

Alternatively, one could specify them in a “Vasicek fashion” as:

(3.28) \[ b_{ii} = k_i, \quad i = 1, 2 \]

After imposing these parametrized volatilities, one may obtain a semiparametric estimate of the FP covariance \( b_{12}(\cdot) \) from (3.8). Constants \( k_i, \ i=1,2 \), must then be obtained beforehand.
The fact that $\sigma_{21} = 0$ implies that the second process will be a true CIR or Vasicek one, so that $k_2$ may be obtained via standard methods.

A natural way to obtain an estimator of the constant in $b_{1t}(\cdot)$ is, first, to consider the kernel estimator of the conditional density using its definition:

\[ (3.29) \quad \hat{f}(x, t \mid y, t') = \frac{\hat{f}(x, y)}{\hat{\pi}(y)} \]

with $x = (x_1, x_2)$ and $y = (y_1, y_2)$, so that the numerator of the right-hand side of (3.29) is the joint density of two observations at the time distance $\tau = (t' + \Delta t) - t' = \Delta t$, which is the interval between observations, and the denominator is the corresponding marginal density. The (product) kernel estimators of the joint and marginal densities are given by, respectively,

\[ (3.30) \quad \hat{f}(x, y) = \frac{1}{nh^3} \sum_{i=1}^{n} K\left(\frac{x_1 - x_{1i}}{h}\right) K\left(\frac{x_2 - x_{2i}}{h}\right) K\left(\frac{y_1 - y_{1i}}{h}\right) K\left(\frac{y_2 - y_{2i}}{h}\right) \]

and

\[ (3.31) \quad \hat{\pi}(y) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{y_1 - y_{1i}}{h}\right) K\left(\frac{y_2 - y_{2i}}{h}\right) \]

where the kernel $K(.)$ is a univariate function such that

(K1) $\int K(u)du = 1$

(K2) $\int u K(u)du = 0$
and the parameter $h$, assumed to be the same for every kernel, is the bandwidth, or smoothing parameter.

The marginal $\hat{f}_i(x_i \mid y)$ of the bivariate conditional $\hat{f}(x \mid y)$ will be:

\[ (3.32) \quad \hat{f}_i(x_i \mid y) = \int \frac{\hat{f}(x,y)}{\hat{\pi}(y)} \, dx_2 = \frac{1}{\hat{\pi}(y)} \int \hat{f}(x,y) \, dx_2 \]

By defining $u = (x_2 - x_{2i})/h$, integrating (3.32) with respect to $u$, and using property (K1) of the kernel function, one obtains

\[ (3.33) \quad \hat{f}_i(x_i \mid y) = \frac{1}{h} \sum_{i=1}^{n} K\left(\frac{x_i - x_{ii}}{h}\right) K\left(\frac{y_{1} - y_{ii}}{h}\right) K\left(\frac{y_{2} - y_{2i}}{h}\right) \]

One should now use (3.33) to compute the marginal variances for selected $y_j = (y_{1j}, y_{2j})$. By defining $v = (x_1 - x_{1i})/h$, integrating with respect to $v$, and using properties (K1) and (K2) of the kernel function, one obtains for the marginal mean:

\[ (3.34) \quad \hat{\mu}_i = \frac{\sum_{i=1}^{n} x_i K\left(\frac{y_{1} - y_{1i}}{h}\right) K\left(\frac{y_{2} - y_{2i}}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{y_{1} - y_{1i}}{h}\right) K\left(\frac{y_{2} - y_{2i}}{h}\right)} \]
and for the marginal variance

\[
\hat{b}_{11} = \frac{\sum_{i=1}^{n} (x_i - \hat{\mu}_i)^2 K \left( \frac{y_{1i} - y_{2i}}{h} \right) K \left( \frac{y_{2} - y_{2i}}{h} \right)}{\sum_{i=1}^{n} K \left( \frac{y_{1i} - y_{2i}}{h} \right) K \left( \frac{y_{2} - y_{2i}}{h} \right)}
\]

(3.35)

Fitting a line through the points \( y_j = (y_{ij}, \hat{b}_{11}(y_j)) \) plus the origin will produce an estimate for \( k_1 \).

Another idea for the estimation of \( k_1 \) may be more demanding and depend on approximations. One may recall that, for small displacements \( \tau = t - t' \), if the derivatives of the FP drifts and volatilities are negligible compared of those of the transition density, the equation to be solved is, approximately:

\[
\frac{\partial}{\partial t} f(x, t \mid y, t') = -\sum_{i=1}^{2} \mu_i(y, \theta) \frac{\partial f(x, t \mid y, t')}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} b_{ij}(y) \frac{\partial^2 f(x, t \mid y, t')}{\partial x_i \partial x_j}
\]

subject to the initial condition

(3.37) \quad f(x, t; y, t') = \delta(x - y)

the solution to the FP equation will be a bivariate normal distribution whose covariance matrix coincides with that of the FP \( b \)'s (see Gardiner (1997), section 3.5):
(3.38) \[ f(x, t \mid y, t') = (2\pi)^{-N/2} \left\{ \det[B(y, t')] \right\}^{1/2} (t-t')^{-1/2} \cdot \exp \left\{ -\frac{1}{2} \left[ x - y - \mu(y, t'; \theta)(t-t') \right]^T [B(y, t')]^{-1} \left[ x - y - \mu(y, t'; \theta)(t-t') \right] \right\} \]

where

\[ B(y, t') = \begin{bmatrix} b_{11}(y, t') & b_{12}(y, t') \\ b_{21}(y, t') & b_{22}(y, t') \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \mu(y, t'; \theta) = \begin{bmatrix} \mu_1(y_1, t'; \theta) \\ \mu_2(y_2, t'; \theta) \end{bmatrix} \]

The equation above represents a Gaussian distribution with covariance matrix \( B(y, t') \) and mean \( y + \mu(y, t'; \theta)(t-t') \).

Recalling that the left hand side of (3.38) is estimable – assuming stationarity – using (3.29)-(3.31), this opens a wide range of possibilities both for estimation and testing. With the parametrizations at stake - in a “CIR / Vasicek fashion” - the parameter \( k_1 \) is not known but, if one estimates parameter \( k_2 \) (eg. by GMM) and takes (3.8) into account, inserting it into (3.38), \( b_{12}(y) \) is eliminated and there is only one unknown in this equation – \( k_1 \). One interesting possibility is to choose a \( k_1 \) which minimizes the Kullback-Leibler discrepancy measure between the associated bivariate normal density and the one estimated nonparametrically. Once this parameter is obtained, one gets \( b_{12}(y) \) in a straightforward manner.

If one considers the assumptions leading to (3.38) reasonable, it is even possible to test a variety of hypotheses comparing the densities \( f(x, t \mid y, t') \) resulting from the two alternatives. For instance, one could compare a variety of parametrizations concerning both the drifts and volatilities of the bivariate process.

4. The Semiparametric Procedure
In order to implement the procedure developed in section 3.2, concerning parametrizations on the diagonal terms of the Itô volatility matrix, we propose to replace the drift parameter vector $\theta$, the densities $\pi(.,.)$, $\pi(.,.)$ and $d\pi_i/dx_i$, $i=1,2$, and the parameters $C_i$, $i=1,2$ by consistent estimators. The densities are estimated using kernel smoothers (see Silverman (1986) for an introduction and Scott (1992) for an advanced treatment), while GMM estimation after discretization of each process yields $\theta$ and $C_i$, $i=1,2$. The only parameter to be estimated is then $\sigma_{12}(.)$, the solution of either (3.15) or (3.22) depending on the assumptions concerning the diagonal terms of the Itô volatility matrix.

If instead one considers the implementation of the procedure suggested in section 3.3, concerning parametrizations on the FP volatilities, we propose to estimate the drift parameter vector $\theta$ and the parameter $k_2$ of the FP volatility of the second process using GMM. The densities $f(.,.), f(.,.), \pi(.)$ and $f_1(.,.)$ should be estimated using kernel smoothers. The first approach suggested concerning the estimation of parameter $k_1$ could be accomplished using OLS estimation. The alternative, based on the density (3.38) deserves more study but, as mentioned before, one interesting possibility could be to choose a $k_1$ which minimizes the Kullback-Leibler discrepancy measure between its associated normal density and the one estimated nonparametrically. Once $k_1$ is obtained, computation of $b_{12}(y)$ is straightforward.

Various regularity conditions on (i) the time series dependence in the data, (ii) the kernel actually used and (iii) the bandwidths, are needed, if asymptotic results are desired for the estimators.

5. An Application to the Brazilian and Argentinian Stock Markets
5.1 The Data

To illustrate our approach, we use daily (logarithmic) returns of the Ibovespa and the Merval, which are, respectively, the main Brazilian and Argentinian stock indexes. The sample is from October 19, 1989 to March 16, 1999, and we considered the index on market closure. It was assumed that Fridays are followed by Mondays, with no adjustment for weekend effects.

Although the series are clearly nonstationary in levels (see figures 1 and 2 in Appendix 1), the returns seem to be stationary (see figures 3 and 4). An interesting feature of the returns is the occurrence of outliers, especially in the Brazilian series, a characteristic of emerging markets. The stationarity assumption was tested for both series (see Table 1), being clearly satisfied. Table 2 shows some summary statistics for the series. One should note that the null hypothesis of normality of the returns is clearly rejected by the Jarque-Bera test (see Davidson and MacKinnon (1993), p. 567), mostly because of kurtosis – this feature will be mentioned again next, when considering the density estimates. As a consequence, estimation methods based on maximum likelihood under the normality assumption are expected to be inefficient.

<table>
<thead>
<tr>
<th>TABLE 1 – Unit Root Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Ibovespa Index</td>
</tr>
<tr>
<td>Merval Index</td>
</tr>
</tbody>
</table>
TABLE 2 – Basic Statistics of the Returns

<table>
<thead>
<tr>
<th></th>
<th>Ibovespa Returns</th>
<th>Merval Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.006694</td>
<td>0.001343</td>
</tr>
<tr>
<td>Median</td>
<td>0.005059</td>
<td>0.000895</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>0.296752</td>
<td>0.039786</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.273208</td>
<td>-1.750782</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>540.7741</td>
<td>72.93948</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normality Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jarque – Bera</td>
<td>27076466</td>
<td>459117.5</td>
</tr>
<tr>
<td>Probability</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

5.2 GMM Estimation

The GMM estimates for univariate Vasicek models were obtained from the following four moment conditions ($\Delta = 1$ day) (see Karlin and Taylor (1981), p. 218, and Ait-Sahalia (1996) for details):

\[(5.1) \quad E f_1(\theta)' = E[ \varepsilon_{t+\Delta}, r_t, \varepsilon_{t+\Delta}, \varepsilon_{t+\Delta}^2 - E[\varepsilon_{t+\Delta}^2 | r_t], r_t (\varepsilon_{t+\Delta}^2 - E[\varepsilon_{t+\Delta}^2 | r_t]) ] = 0 \]

where

\[(5.2) \quad \varepsilon_{t+\Delta} = (r_{t+\Delta} - r_t) - E[(r_{t+\Delta} - r_t) | r_t] \]

\[(5.3) \quad E[(r_{t+\Delta} - r_t) | r_t] = (1 - e^{-\beta \Delta}) (\alpha - r_t) \]

\[(5.4) \quad E[\varepsilon_{t+\Delta}^2 | r_t] = (\sigma^2/2\beta) (1 - e^{-2\beta \Delta}) \]
One should recall that (i) this problem does not reduce to OLS, as we have an overidentified system, (ii) these moments correspond to transitions of length $\Delta$, and are not subject to discretization bias.

The moment conditions above are just a first approximation to the problem. As a matter of fact, ideally, one should also consider the off-diagonal terms of the Itô volatility matrix, as discussed in section 3.

**TABLE 3 – GMM Estimation for the Vasicek Model**

<table>
<thead>
<tr>
<th></th>
<th>Ibovespa Returns</th>
<th>Merval Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$3.893 \times 10^{-3}$</td>
<td>$1.225 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td>(4.10)**</td>
<td>(1.39)*</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$2.410$</td>
<td>$2.794$</td>
</tr>
<tr>
<td></td>
<td>(6.51)**</td>
<td>(3.87)**</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$7.833 \times 10^{-3}$</td>
<td>$3.917 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>(10.81)**</td>
<td>(2.81)**</td>
</tr>
</tbody>
</table>

Notes: (i) The estimates reported are for daily sampling of the returns.

(ii) Heteroskedasticity-robust t statistics are in parentheses.

* Null hypotheses rejected at the 10 percent level of significance.

** Null hypotheses rejected at the 1 percent level of significance.

### 5.3 Density Estimation

The densities of both Ibovespa and Merval returns are characterized by heavy tails, as we illustrate in this section. Consider first the joint Ibovespa and Merval returns density, in
figures 5 and 5a in Appendix 1, and then the nonparametric marginal densities estimates for
Ibovespa and Merval returns (figures 6 and 7) compared to the normal densities with same
mean and variance of each return (figures 8 and 9). As the density estimates are inputs of the
estimation procedure, it is taking account of the heavy tails which characterize the data at
stake, although it is also worth to mention that there are several methods to handle outliers in
a more proper way, which will certainly be done in the future.

5.4 Itô Covariance Estimation: The Double Vasicek and CIR Models

Given the parameter estimates for $\theta_i, C_i$, $i = 1,2$, and the nonparametric density
estimates for $\pi(.,.)$ and $\pi(.,), i=1,2,$ the covariance estimate $\hat{\sigma}_{12}$ solves (3.15) and (3.22),
respectively for the Vasicek and CIR assumptions concerning the elements of the diagonal of
the Itô covariance matrix. It must be pointed out that, somewhat improperly, we shall use in
this exercise the $C_i$ estimates obtained above. This could be avoided by considering moment
conditions which take into account the off-diagonal terms of the Itô volatility matrix, as
already mentioned in section 5.2.

However, one important question concerning the boundary conditions to be used
remains. Consider first the Vasicek model in the case of fixed income. At the point $(x_1,x_2) =
(0,0)$ one may rewrite (3.16)-(3.19) as:

\begin{align}
(5.5) \quad A_1 &= C_2 \pi(0,0) \\
(5.6) \quad A_2 &= \pi_1(0) \\
(5.7) \quad A_3 &= \frac{1}{2} \frac{d\pi_1(x_1)}{dx_1} \bigg|_{x_1=0}
\end{align}
if we impose \( r_j(0) = g_j(0) = \pi(0,0) = 0 \), which is the bivariate counterpart of the boundary condition imposed by Ailt-Sahalia (1996), the result is

\[
(5.9) \quad \sigma_{12}(0) = \left[ -\left( \frac{d\pi_i(x_i)}{dx_i} \right)_{x_i=0} + \left( \sum_{i=1}^3 C_i^0 \frac{d\pi_i(x_i)}{dx_i} \right)_{x_i=0} \right]^{1/2}
\]

It is straightforward to see that one may get a complex-valued boundary condition: imposing \( \sigma_{12}(0) > 0 \) implies that

\[
(5.10) \quad \frac{d\pi_2(x_2)}{dx_2} \bigg|_{x_2=0} < -\frac{C_1^2}{C_2^2} \frac{d\pi_1(x_1)}{dx_1} \bigg|_{x_1=0}
\]

As both \( C_1 \) and \( C_2 \) are assumed to be strictly positive and the derivatives are likely to have the same (positive) sign, the assumption is invalid.

Consider next the Vasicek model in the case of variable income. Assuming that \( \pi_1(x_1^{\text{min}}) = \pi_2(x_2^{\text{min}}) = \pi(x_1^{\text{min}}, x_2^{\text{min}}) = 0 \), one gets, at \( (x_1, x_2) = (x_1^{\text{min}}, x_2^{\text{min}}) \), for (3.16)-(3.19):

\[
(5.11) \quad A_1 = 0
\]
\[
(5.12) \quad A_2 = 0
\]
\[
(5.13) \quad A_3 = \frac{1}{2} \frac{d\pi_i(x_i)}{dx_i} \bigg|_{x_i^{\text{min}}}
\]
leading to:

\[
\sigma_{12}(x_1^{\text{min}}) = \left[ - \left( \frac{d\pi_1(x_1^{\text{min}})}{dx_1} \right) \left( \sum_{i=1}^{2} C_i^2 \frac{d\pi_i(x_1^{\text{min}})}{dx_i} \right) \right]^{1/2}
\]

Once again one may get a complex-valued boundary condition, what in fact happened with the data at stake. Of course, there are several other boundary conditions to be considered but, given the results above, several assumptions were tried, and those actually used were obtained as follows. For the Vasicek model, assume that \(x_1^{\text{min}}\) is such that \(A_1=0\) and \(A_4=0\), a reasonable assumption if someone is considering the data at stake. Under such assumptions, the differential equation (3.15), related to the Vasicek model, may then be written as

\[
A_2 \frac{d\sigma_{12}(x_1^{\text{min}})}{dx_1} \sigma_{12}^{-1}(x_1^{\text{min}}) + A_3 = 0
\]

by straightforward calculations, one gets:

\[
\log(\sigma_{12}(x_1^{\text{min}})) = -\frac{A_3}{A_2} x_1^{\text{min}} + K_2
\]

assuming that \(K_2 = 0\), one gets:
(5.18) \[ \sigma_{12}(x_1^{\min}) = \exp\left\{ -\frac{A_3}{A_2} x_1^{\min} \right\} \]

Alternatively, one may assume that \( \sigma_{12}(x_1^{\min}) = \sigma_{12}'(x_1^{\min}) = w \) and that \( A_1 = 0 \). This allows us to rewrite (3.15) as:

(5.19) \[ (A_2 + A_3) w^2 + A_4 = 0 \]

The solutions are given by:

(5.20) \[ \sigma_{12}(x_1^{\min}) \equiv w = \frac{(-4A_4(A_2 + A_4))^{1/2}}{2(A_2 + A_3)} \]

and

(5.21) \[ \sigma_{12}(x_1^{\min}) \equiv w = \frac{(-4A_4(A_2 + A_3))^{1/2}}{2(A_2 + A_4)} \]

The plots of the solution of (3.15) subject to the boundary conditions (5.18) and (5.20) are shown in Appendix 2. An important remark is that the solutions didn’t converge for a variety of values of \( x_1 \) and \( x_2 \).

Although each boundary condition may originate a completely different behavior of the covariance estimate, one should notice that the shape in each of the two cases considered does not vary too much (see figures in Appendix 2). However, further considerations require the construction of a testing framework, so as to conclude whether variable \( x_2 \) is an argument of the covariance (function) estimates, a question that deserves more study in the future.
6. Conclusions

This thesis has developed a procedure, in an exploratory level, to estimate and test two-dimensional diffusions, with applications to finance. The extension of Aït-Sahalia’s univariate idea to the bivariate case is not immediate, since the solution of the corresponding Fokker-Planck equation can easily become very difficult, if not impossible in analytical terms. By parametrizing the drifts of both processes and imposing restrictions on the terms of the Itô and Fokker-Planck covariance matrices, it is sometimes possible to obtain a nonparametric estimate of the covariance between the processes. However, a delicate issue might still remain, regarding the definition of the boundary conditions for the partial differential equations to be actually solved.

The main message of this thesis is then that extending in a general way Aït-Sahalia (1996)’s framework to a multivariate setting is by no means straightforward. The basic reason is perhaps because the correspondence between the Itô and Fokker-Planck representations in higher dimensions is not the same as in the univariate case. For the methods developed here, the Fokker-Planck equation is more tractable than the Itô representation, suggesting that parametrizations should be made on the latter representation. Notwithstanding, even when imposing sensible assumptions on the drift and volatility functions, one cannot obtain a direct generalization of the univariate method.

Some suggestions for improvement and further research include (i) the development of the testing procedure concerning the differential equations resulting from the parametrizations on the Itô volatilities, already mentioned in section 3.2; (ii) the investigation of the class of diffusions whose densities may be written in closed form, the development of a (nonparametric) testing procedure to identify such class of diffusions, and the implementation of the estimation and testing procedure sketched in section 3.3, including asymptotic results;
(iii) the improvement of the GMM estimation procedure presented in section 5.2; (iv) the
comparison of robust density estimation methods with the standard kernel method used in
section 5.3; (v) alternative boundary conditions for differential equations (3.15) and (3.22), as
already mentioned in section 5.4; (vi) the allowance of different bandwidths and the
abandonment of the product kernel in section 3.3, especially with the availability of more
data.

A particular case that seems to be promising is that of "causal bivariate models" in
which one of the diffusions contributes to the volatility of the other. The appeal of this idea is
immediate – investors could improve their hedging strategies when considering a flexible
estimator of the covariance between two (groups of) assets of a given portfolio, instead of
assuming independence among processes and estimating separately univariate diffusions.

References

Econometrica, 64, 527-560.


Cox, J., J. Ingersoll and S. Ross (1985a): "An Intertemporal General Equilibrium Model of


Oxford: Oxford University Press.


Appendix 1 – Figures

Ibovespa Index

Merval Index
Joint Density: Ibovespa & Merval

Figure 5
Joint Density: Ibovespa & Merval

Figure 5a
Nonparametric Marginal Density: Ibovespa Returns

Nonparametric Marginal Density: Merval Returns
Nonparametric Density: Ibovespa x 'Normal' Ibovespa (in circles)

Nonparametric Density: Merval x 'Normal' Merval (in circles)
Appendix 2 – Semiparametric Covariance Estimation

In this appendix we show the covariance estimates using the alternative boundary conditions developed in section 5.4. First, we plot the covariance function – the solution of \((3.15)\) - subject to the boundary condition \((5.18)\). Then we consider the solution of \((3.15)\) subject to the boundary condition \((5.20)\).

(i) The Solution of \((3.15)\) Subject to Boundary Condition \((5.18)\)

It is worth mentioning that the solution of \((3.15)\) subject to \((5.18)\) did not converge for every value of \(x_1\) (Ibovespa Returns). Thus, our choice was to take the values \(-0.2, -0.1, 0, 0.1, 0.2\) for \(x_2\) (Merval Returns) to perform the numerical solution desired.

An interesting feature of the plot (Figure 10) of the solution is that it may be an indicator of whether variable \(x_2\) is an argument of the covariance (function), although further considerations require the construction of a testing procedure.
Figure 10 - Semiparametric Covariance Estimate: Solution of (3.15) subject to (5.18)
(ii) The Solution of (3.15) Subject to Boundary Condition (5.20)

As in the previous case, the solution of (3.15) subject to (5.20) did not converge for every value of \( x_1 \) (Ibovespa Returns). Thus, our choice was to take the values \(-0.2, -0.1, 0, 0.1\) for \( x_2 \) (Merval Returns) to perform the numerical solution desired.

Just as in the previous case, the plot in Figure 11 may be an indicator of whether variable \( x_2 \) is an argument of the covariance (function), although further considerations require the construction of a testing procedure, a point that deserves more study in the future.
Figure 11 - Semiparametric Covariance Estimate: Solution of (3.15) subject to (5.20)