Universidade de São Paulo

Henning William Menke

Introduction to the Kalman: Applications in Economics

An undergraduate thesis submitted in fulfillment for the degree of Bachelor of Science

Adviser: Pedro Tonelli
IME - USP

January 2012

List of Figures

Contents

Chapter 1

Introduction

1.1 Background

The idea of any filter is to make sense out of noisy data, to make data clearer or separate noise from data. There are several reasons for data to be noisy or unobserved, like measurement errors or frequency of measurement. In the early 1960's this problem was of particular interest to engineers who wished to monitor and control dynamic systems such as a space shuttle launch. It is of extreme importance to have a precise measurement of current conditions (velocity, location etc...) of the shuttle in order to have a successful launch. The modeling method used in dynamic systems is generally called 'State Space Method' and the optimal solution of a linear dynamic system under a gaussian environment is given by the Kalman Filter. State Space modeling may be found in several different subjects such as Control Engineering, Signal Processing, Machine Learning and Time Series Analysis. It's a generic framework used to model a dynamic system, be it in continuous or discrete time, in which one may separate what the user can observe and the real state of the variables of interest in the system. In Control Engineering, one looks at a system and simulates what would happen to it if one changes the control variable, for example what happens to the trajectory of a rocket when fuel injection is activated during flight. In Signal Processing and Time Series Analysis, one tries to extract the state of a system given a measurement of some variables. It may be tracking a boat on the ocean using radar or decomposing the Gross Domestic Product time series into trend, seasonal and level. The notation used in the following text is the same as in Pole *et. al* (1994)[5] and Petris *et. al* (2009)[1]

With the general idea introduced, given a sequence of observations $\mathbf{Y_t}$, of m time series and θ_t as a vector of states, i.e. the variables we are interested in, of dimension p - we can model a dynamic system with an observation equation and a state equation:

$$\begin{aligned} \mathbf{Y_t} &= f(\theta_t, h(v_t)) \\ \theta_t &= g(\theta_{t-1}, u_t, w_t) \end{aligned} \tag{1.1}$$

The first equation is the observation one, which depends on the current state variable and contains some form of measurement error v_t, $p \times 1$ vector, the second is the state equation which, assuming the markovian property, depends on the lagged state, a set of control variables u_t and the innovation w_t, $m \times 1$ vector, which is assumed independent of the measurement error. It does not assume linearity of the variables, the innovations and errors and also doesn't give any structure of their probability distribution - but if assumed linearity and Gaussian errors and innovations, the state space approach is commonly called Dynamic Linear Model(DLM), which is the type of model we will work from now on. Curiously, when the set of states is a subset of the integer set, the state space approach may also be called a Hidden Markov Chain.

1.2 Statistical Basics

In order to understand how the Kalman Filter works, there is a need to develop ideas of conditional probability. The core of Probability theory is to assign a likelihood to all events that might happen under a certain experiment. Kolmogorov's probability axioms state that a probability space is defined by a sample space Ω, an event space E and a probability measure P - but in real life we do not know the real probability measure, i.e. there's uncertainty surrounding any estimate of the real outcome likelihood of events. To help resolve this question, a bayesian approach to probability theory has come into play, which is basically continuously updating our estimate according to new information flow. This approach is called bayesian due to Thomas Bayes and his famous theorem of inverse probability. It states that, let A and B be two events contained in E, the conditional probability of A given that B has occurred is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) * P(A)}{P(B)} \tag{1.2}$$

Another important property we need in dynamic systems is that it is sequential. We know that some variables like the mean height of men in the USA in a given year, the human genome or the position of a car are sequentially ordered and thus allow an online revaluation of our analysis. This means we can conditionally update our estimates each time a new observation arrives or a new measurement is taken. Linking with the bayesian approach above, given a sequence of data or time series $\mathbf{Y_{1:t}}$ and a variable of interest θ_t we can use Bayes' theorem to update the distribution of our variable:

$$P(\theta_t|Y_{1:t}) = \frac{P(Y_{1:t}|\theta_{t-1})P(\theta_t)}{P(Y_{1:t})} \tag{1.3}$$

This means we can compute recursively the distribution of our variable of interest - which in way is exactly what we do in the Kalman filter. It might be a good idea to illustrate it with an example. Say we have θ_t, the height of the water inside a swimming pool and we take daily measurements, $Y_{1:t}$. Let's first consider that θ_t is fixed in time so $\theta_t = \theta$. As our measurement is not precise due to the movement of the waves, there is a difference between Y and the, unobserved, height of the water - an error that we assume to be normally distributed with 0 mean and variance σ^2. Notice that we do not know the value of θ, but we actually want to estimate it from our observations. Mathematically speaking we have:

$$Y_t = \theta + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{1.4}$$

So $Y_t \sim \mathcal{N}(\theta, \sigma^2)$ and if we assume that our initial estimate of $\theta_0 \sim \mathcal{N}(m_0, C_0)$ is normally distributed with mean m_0 and variance C_0, we can use our bayesian algorithm to update our estimate of θ when a new observation is made. Remembering that constants in distribution may be drop as we can always normalize our distribution afterwards, we have the updated distribution of θ:

$$\begin{aligned} P(\theta|Y_1) &\propto exp\left\{-\frac{(Y_1-\theta)^2}{2\sigma^2}\right\} exp\left\{-\frac{(\theta-m_0)^2}{2C_0}\right\} \\ &\propto exp\left\{\tfrac{1}{2\sigma^2 C_0}\left[(\sigma^2+C_0)\theta^2 - 2(\sigma^2 m_0 + C_0 Y_1)\theta\right]\right\} \end{aligned} \tag{1.5}$$

Using the fact that $X \sim \mathcal{N}(\mu, \sigma^2) \propto exp\left\{-\frac{X^2-2\mu X}{2\sigma^2}\right\}$ we may reach to the conclusion that the new distribution of θ given this observation Y_1 is $\theta \sim \mathcal{N}(m_1, C_1)$ where m_1 is $\frac{C_0}{C_0+\sigma^2}Y_1 + \frac{\sigma^2}{C_0+\sigma^2}m_0$ whereas C_1 is $\frac{C_0\sigma^2}{C_0+\sigma^2}$. Extending to further observations 1,...,n we would have then:

$$P(\theta|Y_{1:n}) \propto \prod_{t=1}^{n} exp\left\{-\frac{(Y_t-\theta)^2}{2\sigma^2}\right\} exp\left\{-\frac{(\theta-m_0)^2}{2C_0}\right\}$$
$$\propto exp\left\{\frac{1}{2\sigma^2 C_0}\left[(\sigma^2+nC_0)\theta^2 - 2(\sigma^2 m_0 + nC_0\bar{Y})\theta\right]\right\}$$

Then we would get $\theta|Y_{1:n}$ is $\theta \sim \mathcal{N}(m_n, C_n)$ where m_n is $\frac{C_0}{C_0+\sigma^2/n}\bar{Y} + \frac{\sigma^2/n}{C_0+\sigma^2/n}m_0$, whereas C_n is $\frac{C_0\sigma^2}{nC_0+\sigma^2}$. Of course, we could also do it *recursively*, i.e. calculating $P(\theta|Y_{1:n})$ using $P(\theta|Y_{1:n-1})$ and the new observation Y_n. Using the same methodology we reach:

$$\mathrm{m}_n = \frac{C_{n-1}}{C_{n-1}+\sigma^2}Y_{n-1} + \frac{\sigma^2}{C_{n-1}+\sigma^2}m_{n-1}$$
$$= m_{n-1} + \frac{C_{n-1}}{C_{n-1}+\sigma^2}(Y_n - m_{n-1})$$
$$C_n = \frac{C_{n-1}\sigma^2}{C_{n-1}+\sigma^2}$$

Interpreting these equations show that our estimate of the height of the water pool is a weighted average of our initial estimate and mean of our measurements, while the weights are proportional to the noise of one another, i.e. the higher the noise in our initial estimate the more we will trust our measurements whereas the higher the noise in our the measurement errors is, we take longer to fully incorporate the observations we have made. Interesting to see is that in the recursive equation we can interpret our updated estimate as our prior estimate plus the error in our prediction weighted by what we already will call *Kalman gain*.

1.3 Dynamic Linear Models

A general DLM is a linear recursive system that takes such form:

$$\begin{aligned} \mathbf{Y_t} &= F_t'\theta_t + w_t, \quad w_t \sim \mathcal{N}(0, V_t) \\ \theta_t &= G_t\theta_{t-1} + v_t, \quad v_t \sim \mathcal{N}(0, W_t) \end{aligned} \tag{1.6}$$

Where F_t is a $m \times p$ matrix and G_t is a $m \times m$ matrix. W_t is the $m \times m$ covariance matrix of the innovations and V_t is the $p \times p$ covariance matrix of the measurement errors. A particular case is the local level model or random walk with noise, which aproximates the series as local means. See figure 1.1 for a visualization of what is meant.

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma^2_\epsilon) \\ \mu_t &= \mu_{t-1} + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \sigma^2_\zeta) \end{aligned} \tag{1.7}$$

A good example of an application of the local level model is the level of a river[1]. In this case we estimate a model for the level of the Nile river during a 2 centuries period. In the figure we can notice that the river has two distinct levels during different periods. Between 1871 and 1900, the river oscillates around the 1100 level whereas in the 1900 to 1970 period the river is consistently under the 1000 level.

Further assumption is made about the distribution, the covariances are time invariant. In this case, the matrixes F and G are time invariant and equal to a scalar 1. If $\sigma^2_\zeta = 0$, $\mu_t = \mu$, μ is equal to $\overline{y}$, a constant mean. Another special case is the linear growth model, or local level with drift or trend:

[1]This example is located in Petris *et. al* (2009)[1], pages 7 and 53-58

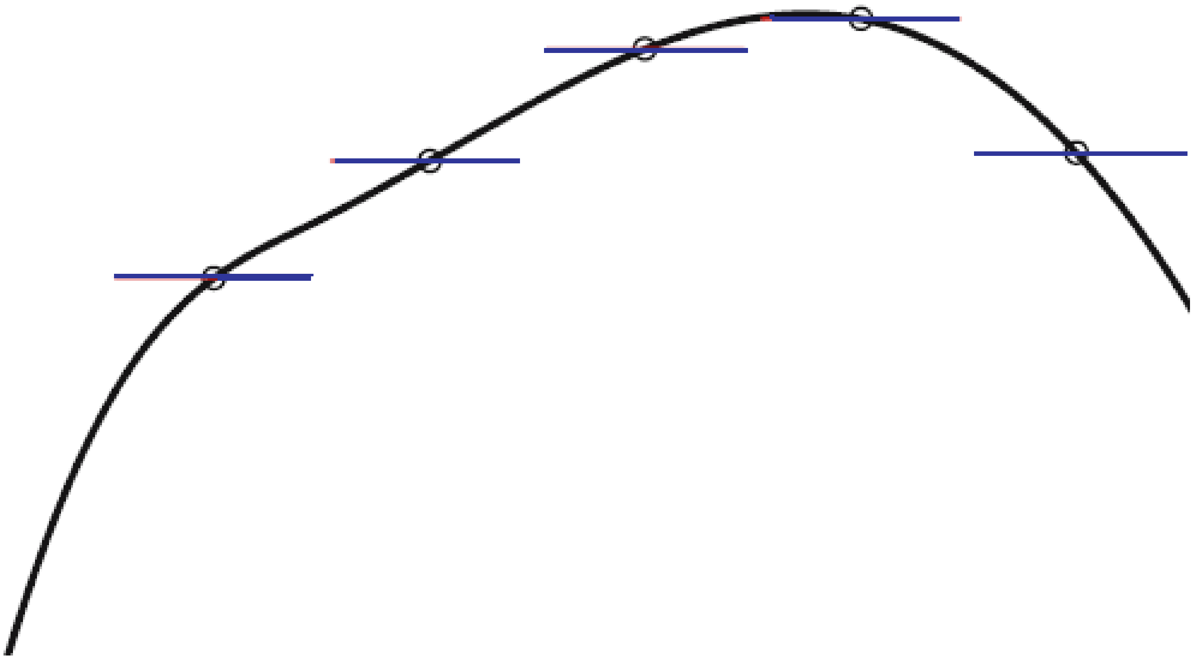

FIGURE 1.1: Example of a series and the local level model

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, & \epsilon_t &\sim \mathcal{N}(0, \sigma^2_\epsilon) \\ \mu_t &= \mu_{t-1} + \theta_t + \zeta_t, & \zeta_t &\sim \mathcal{N}(0, \sigma^2_\zeta) \\ \theta_t &= \theta_{t-1} + \nu_t, & \nu_t &\sim \mathcal{N}(0, \sigma^2_\nu) \end{aligned} \tag{1.8}$$

See figure 1.3 for a linear growth model example and also figure 1.4 for a real life example where such a model could be used - the world's internet usage over time. In another terminology, it could be said that the series contains a stochastic level and trend. In this case, the matrixes would look like:

$$\theta_t = \begin{bmatrix} \mu_t \\ \theta_t \end{bmatrix} \; F = \begin{bmatrix} 1 \\ 0 \end{bmatrix} G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} W = \begin{bmatrix} \sigma^2_\zeta & 0 \\ 0 & \sigma^2_\nu \end{bmatrix}$$

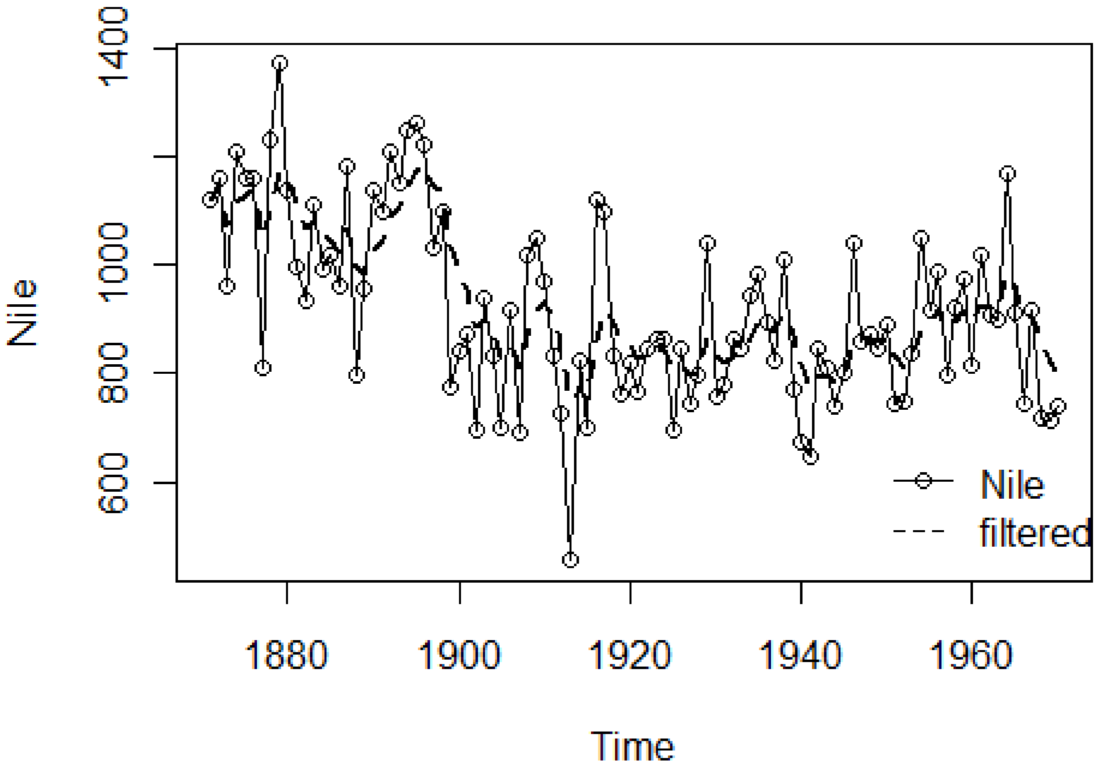

FIGURE 1.2: Nile river

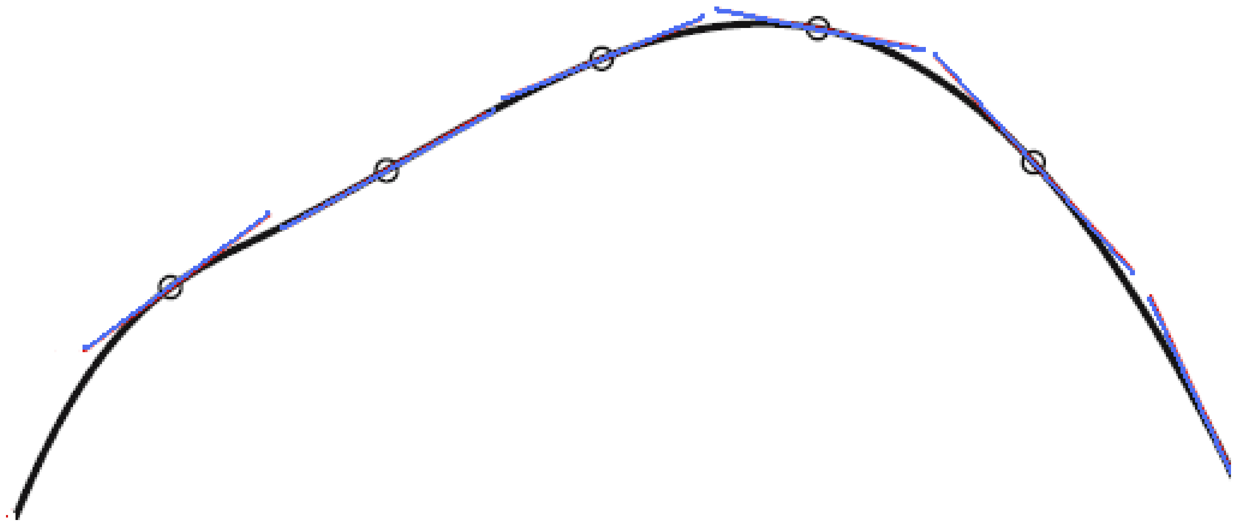

FIGURE 1.3: Example of a linear growth model

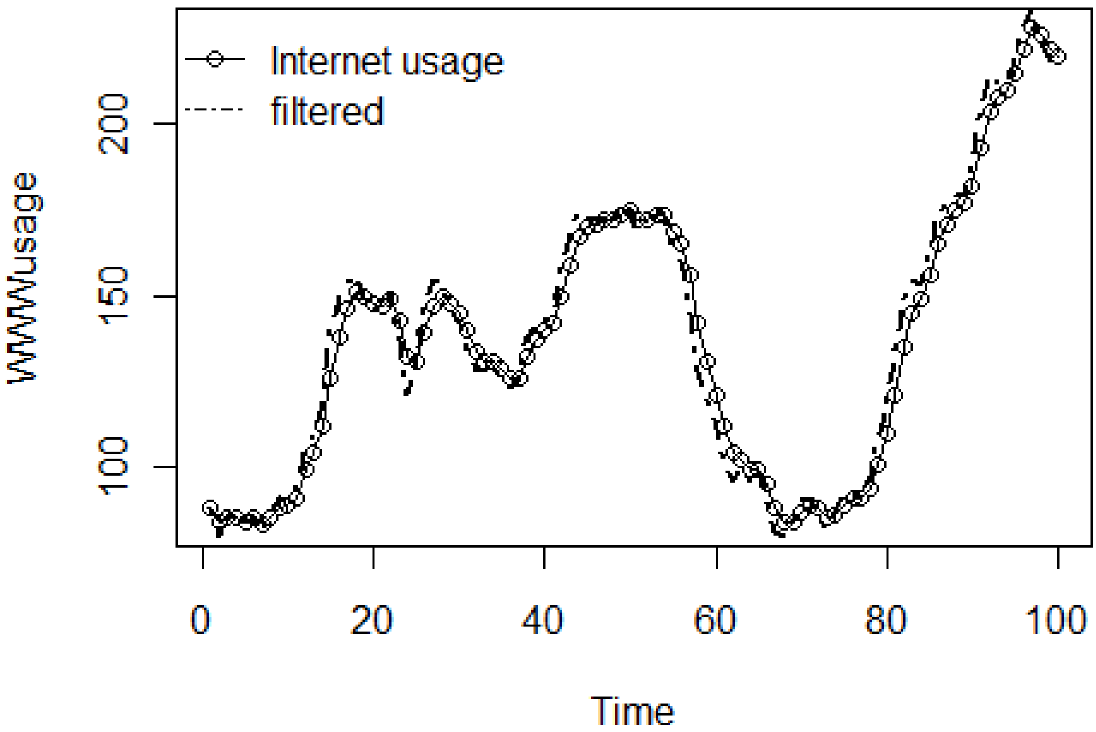

FIGURE 1.4: World's internet usage

Another model to be shown is the dynamic regression model, in which the regression coefficient is allowed to vary in time.

$$\begin{aligned}
y_t &= \mu_t + \beta_t x_t + \epsilon_t, \qquad & \epsilon_t \sim \mathcal{N}(0, \sigma^2_\epsilon) \\
\mu_t &= \mu_{t-1} + \theta_t + \zeta_t, \qquad & \zeta_t \sim \mathcal{N}(0, \sigma^2_\zeta) \\
\theta_t &= \theta_{t-1} + \nu_t, \qquad & \nu_t \sim \mathcal{N}(0, \sigma^2_\nu) \\
\beta_t &= \beta_{t-1} + \xi_t, \qquad & \xi_t \sim \mathcal{N}(0, \sigma^2_\xi)
\end{aligned} \tag{1.9}$$

$$\theta_t = \begin{bmatrix} \mu_t \\ \theta_t \\ \beta_t \end{bmatrix} \; F = \begin{bmatrix} 1 \\ 0 \\ x_t \end{bmatrix} \; G = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \; W = \begin{bmatrix} \sigma_\zeta^2 & 0 & 0 \\ 0 & \sigma_\nu^2 & 0 \\ 0 & 0 & \sigma_\nu^2 \end{bmatrix}$$

Interesting to note is that if $W = 0$, we return to the classic linear regression model. There are several other models including seasonality and multivariate settings. A particular model from the regression one is the auto-regressive model In general such model can be described as:

$$Y_t = \sum_{i=1}^{n} \phi_i Y_{t-i} + \epsilon_t$$

We can put it on state space form if we choose $\theta = [\phi_1 \dots \phi_n]$ and $F = [Y_{t-1} \dots Y_{t-n}]$. This shows how general the state space modeling approach really is.

1.4 Filtering

In the DLM case, the Kalman filter is the optimal filter[2], in the mean squared error sense, to estimate the value of the state at a given point in time t, considering the observations up to t, i.e. using only the values of $\{y_i\}_{i=0...t}$. It means that given the correct state space structure, and the correct variances of the state innovations and measurement errors, the Kalman Filter will minimize $||\mathbf{Y_t} - \widehat{\mathbf{Y}}_{(\mathbf{t-1}):(\mathbf{t})}||_2$[2], delivering the best one step ahead point predictor. Remembering that the assumption of gaussian erros, we must only estimate the mean and variance parameters to fully characterize the distribution of the states. The recursive algorithm of the Kalman filter, with $\theta_0 \sim N(m_0, C_0)$, following Petris *et. al* (2009)[1], is now presented. First we calculate the parameters of the predictive distribution of θ_t given $y_{1:t-1}$:

$$\begin{aligned} a_t &= E\left(\theta_t \left| y_{1:t-1}\right.\right) = G_t m_{t-1} \\ R_t &= Var\left(\theta_t \left| y_{1:t-1}\right.\right) = G_t C_{t-1} G_t' + W_t \end{aligned} \tag{1.10}$$

Then we must calculate the parameters of the predictive distribution of y_t given $y_{1:t-1}$:

$$\begin{aligned} f_t &= E\left(y_t \left| y_{1:t-1}\right.\right) = F_t' a_t \\ Q_t &= Var\left(y_t \left| y_{1:t-1}\right.\right) = F_t' R_t F_t + V_t \end{aligned} \tag{1.11}$$

[2]The prediction error of t at t-1

Before we re-update our estimates of mean and variance, we need to calculate the prediction error and what is called the Kalman gain:

$$\begin{aligned} e_t &= y_t - f_t \\ K_t &= R_t F_t Q_t^{-1} \end{aligned} \tag{1.12}$$

Finally, the filter gives out the estimates for the prediction in t+1:

$$\begin{aligned} m_t &= a_t - K_t e_t \\ C_t &= R_t - K_t Q_t K_t' \end{aligned} \tag{1.13}$$

Where:

$$\begin{cases} a_t & \text{is the prediction of the state for t} \\ m_t & \text{is the expected value of the state at time t} \\ f_t & \text{is the forecast of the observation at time t} \\ e_t & \text{is forecast error} \\ R_t & \text{is signal variance} \\ Q_t & \text{is the noise variance} \\ K_t & \text{is the Kalman gain, a ponderation of the signal and noise variance} \end{cases}$$

The most important step is the calculation of the Kalman gain. As the we can inspect from $m_t = a_t - K_t e_t$, the predicted state for t+1 is equal to a weighted average of the predicted state at t and the prediction error of t, where the weight is known as Kalman gain - it is how you balance what you already know about your variables of interests and the new information and evidence you gain from the observations. The derivation of the Kalman gain is given by:

1.5 Deriving the Kalman gain

The Kalman filter is derived by minimizing the mean squared error and thus discovering the Kalman gain formula. The error we are referring to is $\theta_t - m_t$. Remembering that $C_t = E[(\theta_t - m_t)(\theta_t - m_t)']$, $R = E[(\theta_t - a_t)(\theta_t - a_t)']$ and $m_t = a_t + K_t(y_t - m_t)$, we propose that $K = RF'Q^{-1}$:

Minimizing the mean squared error is equal to minimizing the trace of C_t, which can be rewritten in the following form:

$$C_t = E[(\theta_t - a_t + K_t(Y_t - m_t))(\theta_t - a_t + K_t(Y_t - m_t)']$$

$$C_t = E[(I - K_tF)(\theta_t - a_t) - K_tv_t)((I - K_tF)(\theta_t - a_t) - K_tv_t)']$$

Assuming the model $Y_t = F'\theta_t + v_t$ is correct and that the measurement error v_t is uncorrelated with the state and its estimate, we have:

$$C_t = (I - K_tF)E[(\theta_t - a_t)(\theta_t - a_t)'](I - K_tF)' + K_tV_tK_t'$$

$$= R_t - K_tFR_t - R_tF'K_t' + K_t(FR_tF + V_t)K_t' = R_t - K_tFR_t - R_tF'K_t' + K_tQ_tK_t'$$

The trace of $C_t - T(C)_t$ is: $T(R_t) - 2T(K_tFR_t) + T(K_tQ_tK_t')$. The first order condition of this problem is:

$$\frac{dT(C_t)}{dK_t} = 0 - 2FR_t' + 2K_tQ_t = 0 \Rightarrow K_t = RF'Q^{-1}$$

1.6 Estimation

Of course, in the real world, there's no certainty if the model used is correct, nor do we have the the variances W_t and V_t, they are unobservable. In order to estimate these variances, also known as *hyperparameters*, one may use maximum likelihood estimation, i.e. run the Kalman Filter several times until one finds the *hyperparameter* vector that maximizes the likelihood of the sample $\{y_i\}_{i=0...T}$ have come from the estimated probability distribution - i.e. the likelihood of the decomposed forecast errors. Another way of estimating the *hyperparameters* is to use sampling methods like Gibbs Sampler or Metropolis-Hastings Algorithm, or more in general Monte Carlo Markov Chain methods, which after an initial estimate of the *hyperparameters*, the method samples a new estimate using a proper probability function.

Chapter 2

Applications

2.1 Introduction

Now we apply the concepts we've learned in the previous chapter to the Brazilian economy. Three different applications have been chosen to show how flexible and interesting this way of modeling time series that is. The applications are focused on macroeconomic and financial time series.

The first application is the estimation of the output gap time series of Brazilian economy. Economists generally look into GDP - Gross Domestic Product - in order to diagnose the current state of the economy and through years of research, they have developed the concepts of business cycles and potential GDP. Potential GDP is the product that is achievable by a country if all resources, labour and capital, are efficiently and sustainably allocated. The business cycle is more of an evidence or empirical fact that economies endure periods of great prosperity and recessions and thus creating a cycle of boom and bust. The output gap is then a measure of the current business cycle, i.e. the difference between potential and current GDP.

The second application is to model the brazilian exchange rate. The exchange rate is a measure of what the market expects in order to exchange one currency for another, i.e. how many units of a foreign currency is required in order to acquire one unit of the domestic currency. This rates has several implications for a country such as buying and selling goods and services with the rest of the world might be cheaper or more expensive depending on how the current rate is. An exporter that produces goods that are priced using a foreign currency wants a devalued currency. As his costs are based on the domestic currency, the same revenue will yield a bigger profit whether the currency is under or overvalued. The case for the importer or domestic consumer is the reciprocal as their income is denominated in domestic currency, but goods and services they might buy are based on the foreign currency. The question then becomes, when is an exchange rate over or undervalued? Using several macroeconomic variables and our modeling method we try to answer that question.

The last application is related to the Brazilian financial market. In finance theory, there is an elegant model called CAPM - Capital asset pricing model - which states that the expected excess return of a given risky asset should be equivalent to the return of the risk-free asset plus the risk premium times an adjustment factor known as β. The idea is to estimate this β - which can be a time-varying factor - for several stocks of the main Brazilian stock index, Ibovespa.

2.2 Modeling the output gap

The GDP is a measure of what a country produces and can be calculated by the final demand of goods and services and often is expressed as a simple sum of private and government consumption and investments. The potential GDP is the sustainable production output, or final demand, that a country can achieve without putting too much pressure on domestic prices or generates too much unemployment - it can be viewed as an equilibrium output. Without going too deeply into economic theory, external shocks like oil prices, liquidity crisis, geopolitical issues etc... can create distortions in the economy and thus deviate the current GDP path from potential GDP. These deviations then form what is known as output gap or the business cycle.

How can we model these unobservable variables? First let's assume that as population and capital stock grow, potential GDP also has a tendency to grow - so we can use a linear growth model for it. For the business cycle, if we assume that potential GDP acts as an 'attractor', i.e. current GDP tends to approach the potential GDP, we can model it as mean-reversion process using an AR model. Noticing that current GDP can then be decomposed as a sum of potential GDP and business cycle, we have:

$$
\begin{aligned}
&GDP_t = GDP_t^{Pot} + BC_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2) \\
&GDP_t^{Pot} = GDP_{t-1}^{Pot} + \Delta_t + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \sigma_\zeta^2) \\
&\Delta_t = \Delta_{t-1} + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \sigma_\zeta^2) \\
&BC_t = \phi_1 BC_{t-1} + \phi_2 BC_{t-2} + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, \sigma_\nu^2)
\end{aligned}
\tag{2.1}
$$

Now that we have a model, we look a little bit closer at the data we will manipulate. We use IBGE's - the brazilian institute of geography and statistics - quarterly seasonalizy adjusted GDP series measured in chained 1995 reais from 1975 to 2011. In order to create a more stable model we use the log transformation. The figures show how Brazil has grown and how the growth rates have varied through time.

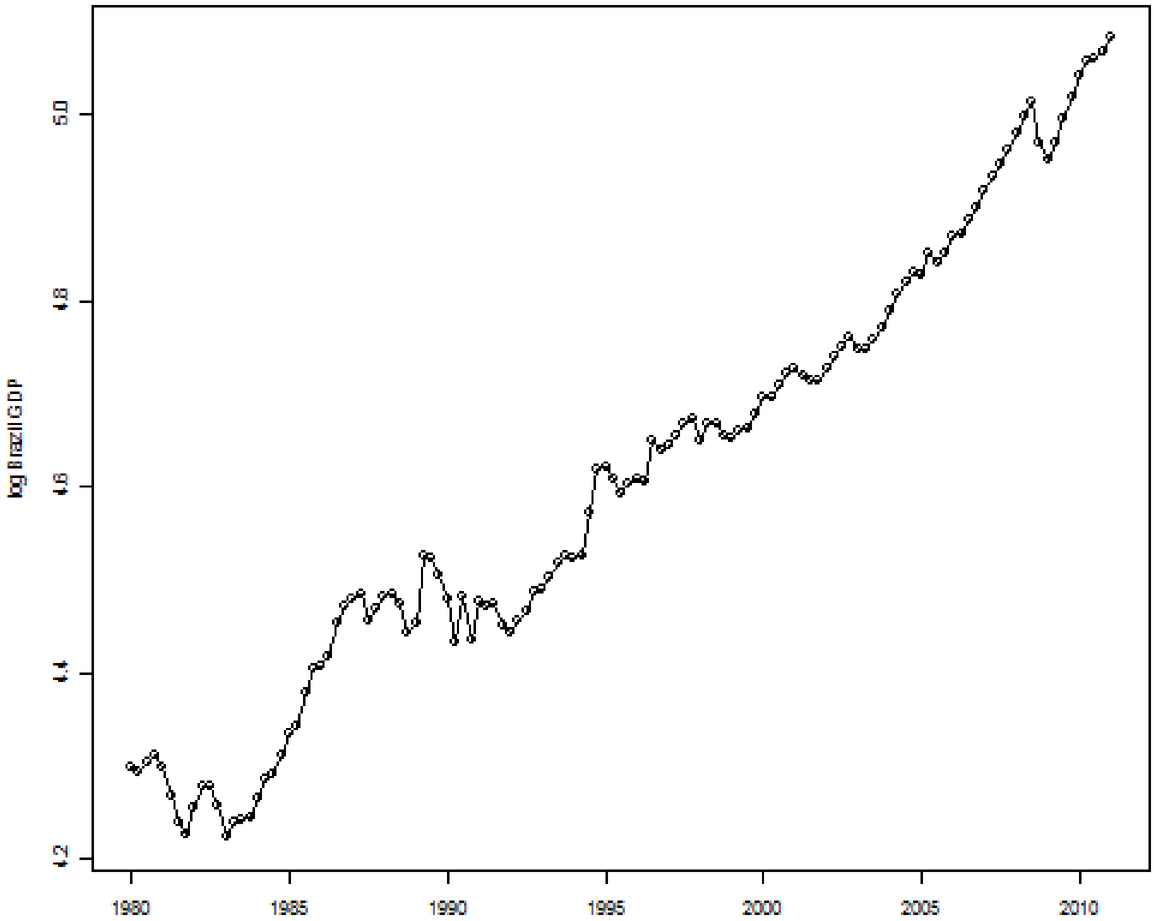

Figure 2.1: Brazilian GDP

One way to analyze the output of our model is to visualize figure 2.3 and draw some parallels between historic moments of Brazilian and the business cycle. As we can see, in the early 1980's real GDP was much lower than potential GDP whereas after 1985 real GDP was significantly higher than

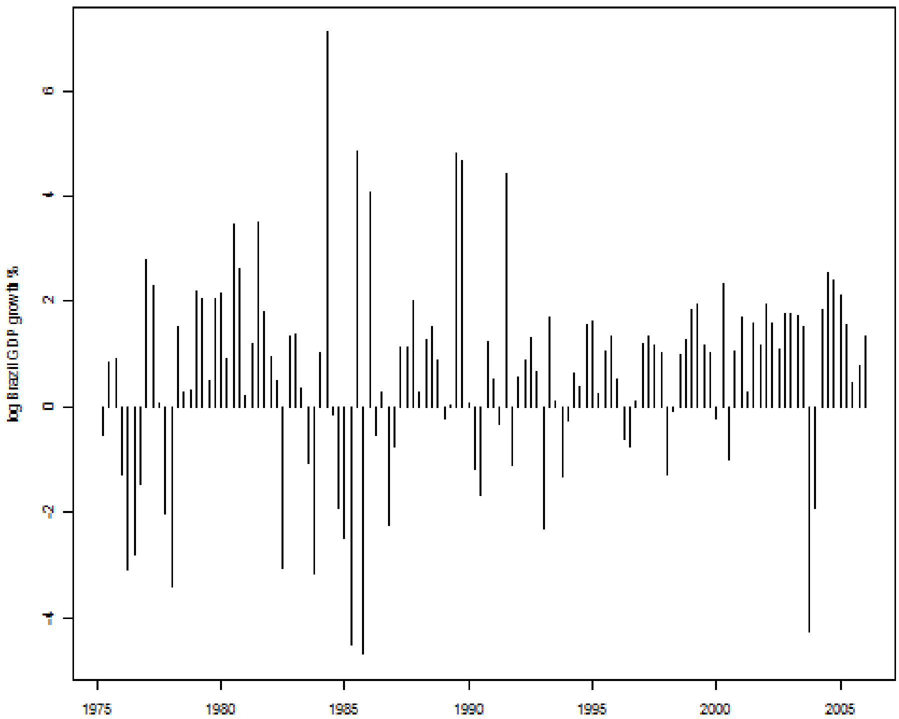

FIGURE 2.2: Brazilian GDP

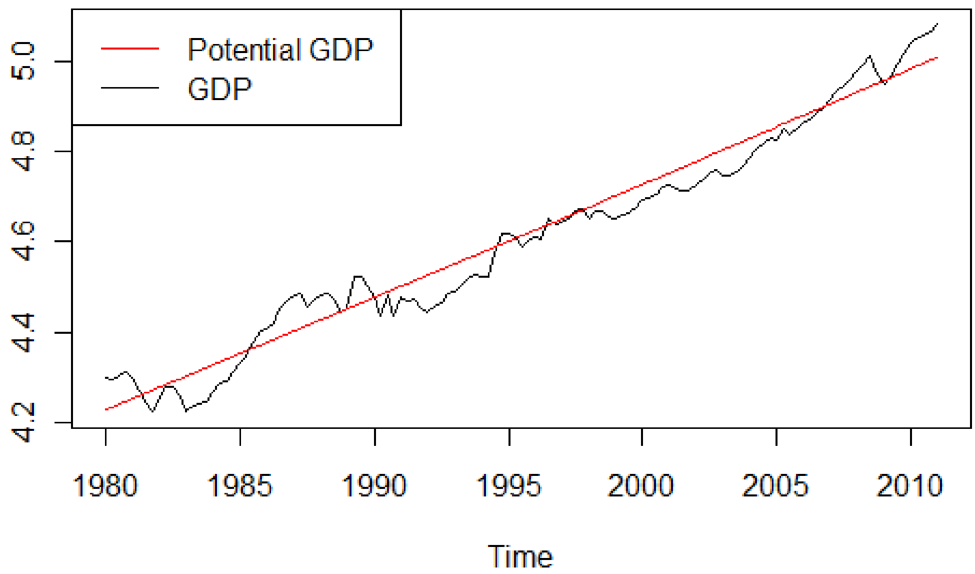

FIGURE 2.3: The time series of real GDP and potential GDP

potential GDP. An educated guess to answer the potential question of why this deviation from potential GDP occurred is to check history of the world economy and the brazilian politics. In 1979, the 2nd oil crisis ensued which not only lowered world GDP, but also ignited the Latin American debt crisis, where several countries defaulted in their debt, including Brazil in 1983. And why did real GDP bounce back and grow even beyond potential? In 1985 the military regime gave place to a new democracy, which in order to bring legitimacy into the government, it boosted growth so that the population would not try to bring the old order back. During the 90's real GDP showed a mild under-performance in relation to potential GDP, but nothing out of the ordinary. Another interesting period is post 2004, in which, again, real GDP overshoots potential GDP and coincides with the change in political dominance from right to left, or in party terms, from PSDB to PT.

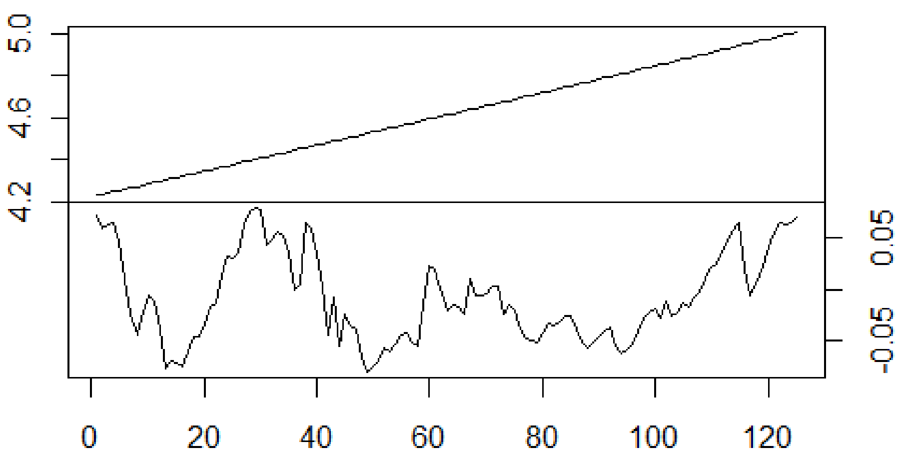

FIGURE 2.4: The decomposition between potential GDP and the business cycle

2.3 Modeling the Brazilian Real

The Brazilian Real was adopted as the common brazilian currency during the Real plan in 1994, whose first objective was to fight and drive down inflation rates. Until 1999, the Real worked under a crawling peg regime, i.e. the central bank slowly depreciated the currency in order to keep pace with the inflation rate differential of United States and Brazil. The regime worked very well to bring inflation to more normal levels than the ones seen during 80's, but during the stress of financial markets of 1997 and 1998 due to the Asian and Russian crisis, ultimately the central bank did not have any reserves left to maintain credible control of the exchange rate and thus the regime change to a floating one.

This little introduction is a caveat for our modeling of the exchange rate. In this application we will use the dynamic linear regression to explain the movements of the Brazilian Real. A normal regression would leave the regression coefficients static over time, but in a dynamic one the coefficients vary over time which would allow more information to flow into the model and give a better explanation of how these variables interact.

We will use the following variables to explain the exchange rate path after the change in currency regime, circa 2000:

DXY - It's an index of the value of the United States dollar relative to a basket of foreign currencies. This variable is important as it is a measure of broad USD weakness or strength.

CRB - A commodity index as Brazil is a big commodities producer and exporter.

One year brazilian interest rates differential - Interest rates play an important part of asset allocation and as such, a higher brazilian interest rate relative to US would increase the propensity of an investor allocating more money into Brazil and thus increasing demand for the Brazilian Real. A proxy used in this paper is the difference between the 1 year PreDI swap rate and the US 1 year.

The justification for these variables is due to common sense and economic theory. As in Marcal and Barbieri (2010) [3] they cite four main fundamental determinants for the exchange rate: Terms of Trade - the price of the country's exports relative to the price of imports, Net Foreign Asset Position, Productivity difference between tradable and non-tradable sectors and Interest rate Parity. In order to have a significant amount of data we'd like to use daily observations and as such we use some available proxies for these variables - for the Terms of Trade we use the CRB, as our exports is mostly made out of commodities. The DXY index is used in order to gauge other factors affecting the USD but are not Brazil related, but still influence the brazilian exchange rate.

The data used can be visualized through the following figures and table:

	IR	BRL	DXY	CRB
Mean	13.31	2.25	90.8	0.32
Standard Deviation	5.287214	0.530506	13.71959	0.088695
$\rho(1)$	0.998485	0.99859	0.999335	0.999838

The model which we will estimate is the following:

FIGURE 2.5: The BRL and the DXY

FIGURE 2.6: The BRL and the interest rate differential

$$\begin{aligned} & BRL_t = \alpha_t + \beta_{1,t} CRB_t + \beta_{2,t} DXY_t + \beta_{3,t} Diff_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2) \\ & \beta_{i,t} = \beta_{i,t-1} + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \sigma_\zeta^2) \end{aligned} \tag{2.2}$$

The results are shown in the next figures as we analyze the dynamic of each factor.

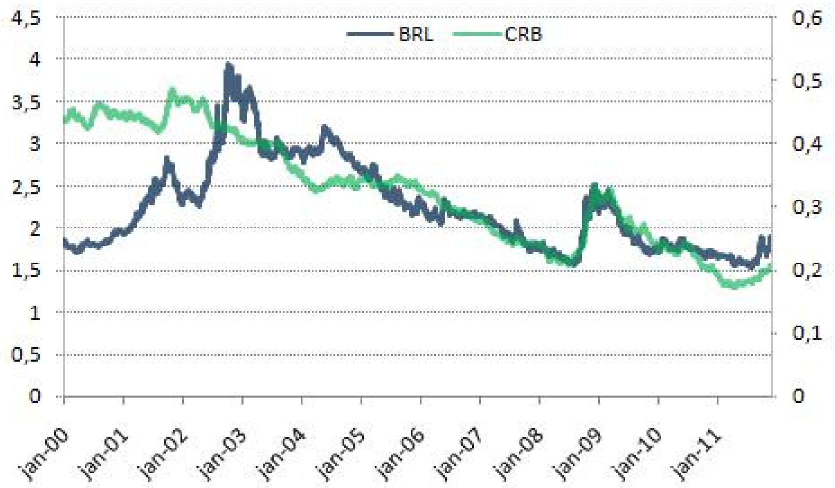

FIGURE 2.7: The BRL and the Commodity Index (inverted scale)

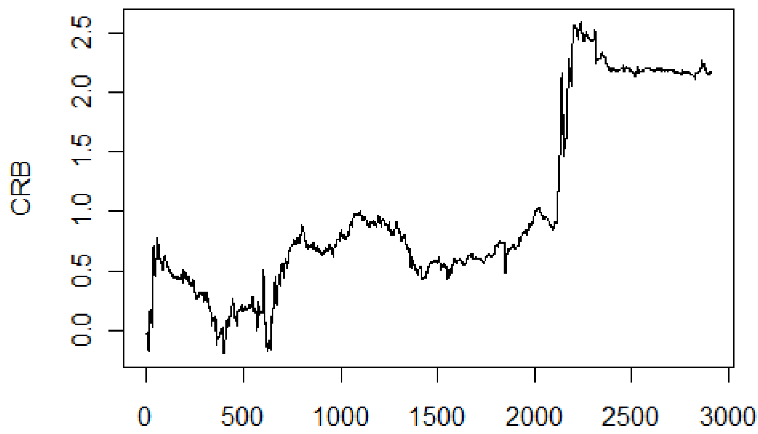

FIGURE 2.8: The evolution of CRB's β

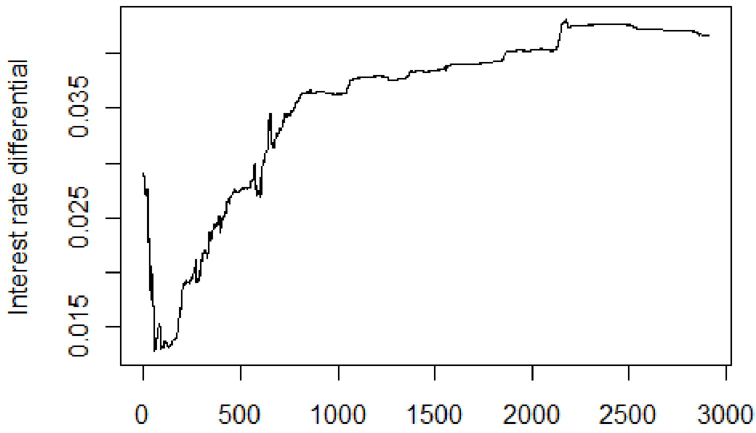

Figure 2.9: The evolution of Interest Rate Differential's β

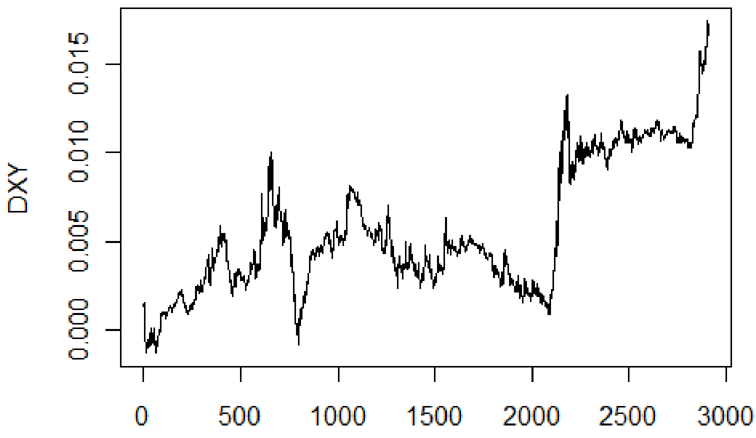

Figure 2.10: The evolution of DXY's β

The most interesting thing is the fact that after observation 2000 (circa 2008) we have a major change in DXY's and CRB's $\beta's$ increasing the weight on these variables on the exchange rate, which most likely is linked to the 2008 financial crisis and it's consequences to global risk assets.

2.4 Dynamic CAPM using Brazilian Stocks

CAPM, or Capital Asset Pricing model, is an important financial theory first proposed by William Sharpe (1964)[6] after the work of optimal portfolio theory of Harry Markowitz (1952)[4]. CAPM, under some fairly constraint conditions, tries to explain the behavior of asset returns in equilibrium conditions, i.e. in a equilibrium state, given asset risks and investors preferences, how much should an asset return for a certain level of risk. Important to note is that the risk being accounted for is the systematic risk, as the remaining idiosyncratic risk can be diversified away by adding more assets to the portfolio. The systematic risk can also be named market risk, i.e. the risk that impacts all assets - even though not equally. CAPM in a nutshell is a linear regression model that relates linearly an asset (a stock in this case) to the overall market return, or in mathematical terms:

$$R_{i,t} = \alpha_t + \beta_1 R_t^M + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

(2.3)

Where $R_{i,t}$ is asset i's return at time t and R_t^M is the markets return at time t, while the regression coefficient β has an intuitive meaning, it is the sensitivity of asset i to systematic risk. CAPM was build as a two-period model and most implementations assume that β is fixed in time. Our application is to check β's time homogeneity is true in the brazilian market for two of the main stocks: Vale do Rio Doce and Petrobras. The data used is the closing prices of each stock from 1/1/2000 to 1/12/2011 and the following chart shows the evolution of each stock price during that period. For the market return we use the Ibovespa index for the same period.

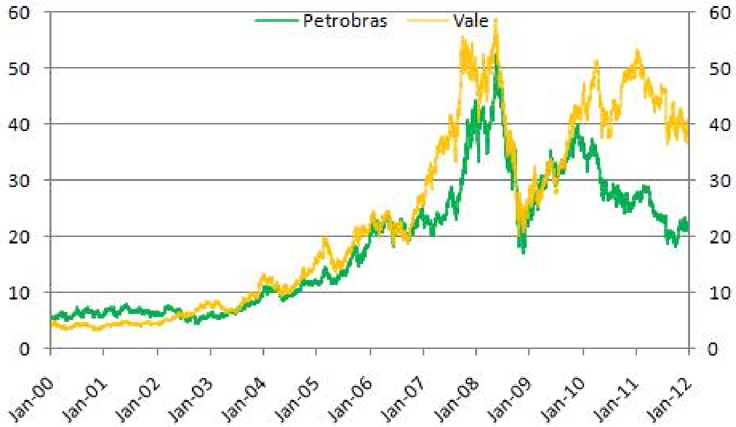

FIGURE 2.11: Price of Vale and Petrobras

FIGURE 2.12: Ibovespa Index

The results were interesting, as the figures show the β's of each stock are not constant, even though the average of each β is very similar to simple linear regression coefficient as shown in the table below.

	Petrobras	Vale
Mean	0.86	0.82
Std Dev	0.177	0.4
LM Coefficient	0.89	0.84

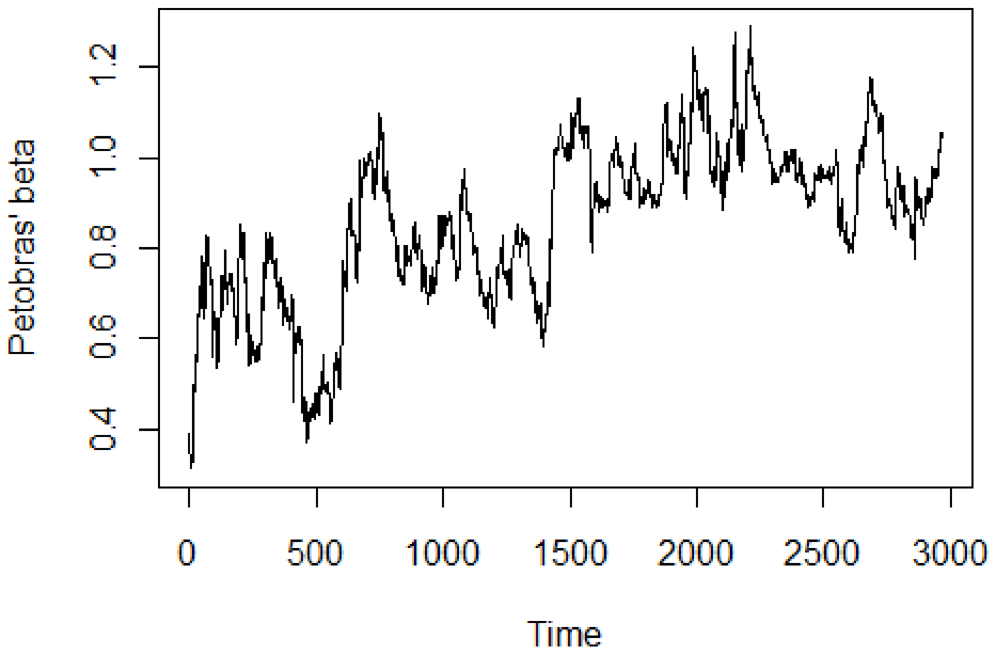

FIGURE 2.13: Evolution of Petrobras' β

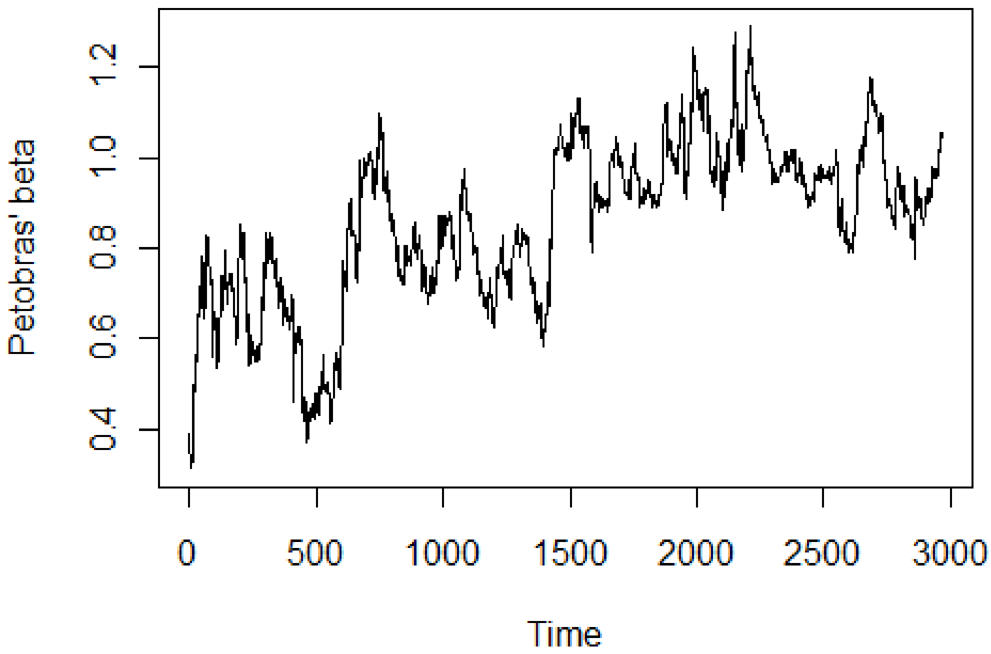

FIGURE 2.14: Evolution of Vale's β

2.5 Conclusions

This study has been conducted in order to show the flexibility and potential of using dynamic linear models in real-life examples. Even though the examples here shown are only relevant to the realm of economics, there are several other applications to engineering, biology, physics that are made possible or easier using the same or similar methodology used in the previous pages such as Kalman Filter. The three examples use the same principles, whereas the models estimated are very different from each other.

Bibliography

[1] Campagnoli, P., Petrone, S., & Petris, G. 2009. *Dynamic Linear Models*. Springer.

[2] Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *J. Basic Engineering, Transactions ASMA Series D*, 35–45.

[3] Marcal, E., & Barbieri, F. 2010. "Quo Vadis Real? Estimating the Brazilian Real Exchange Rate Misalignment in Vector Error Correction Model with Structural Change". *Working Paper Series, FGV*.

[4] Markowitz, H. 1952. "Portfolio Selection". *Journal of Finance*, **7**(1), 77–91.

[5] Pole, A., West, M., & Harrison, J. 1994. Applied Bayesian Forecasting.

[6] Sharpe, W. 1964. "Capital asset prices: A theory of market equilibrium under conditions of risk". *Journal of Finance*, **19**(3), 425–442.