

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

Análise Estatística do Modelo de Nelson e Siegel

Marcelo Bertini Brocco

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

Análise Estatística do Modelo de Nelson e Siegel

Marcelo Bertini Brocco

Orientador: Prof. Dr. Márcio Alves Diniz

São Carlos/SP

2013

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

Análise Estatística do Modelo de Nelson e Siegel

Dissetação apresentada ao Departamento de Estatística da Universidade Federal de São Carlos - DEs/UFSCar, para obtenção do Título de Mestre em Estatística.

São Carlos/SP

2013

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

B863ae

Brocco, Marcelo Bertini.

Análise estatística do modelo de Nelson e Siegel /
Marcelo Bertini Brocco. -- São Carlos : UFSCar, 2013.
96 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2013.

1. Estatística. 2. Inferência bayesiana. 3. Taxas de juros.
4. Métodos de estimação. 5. Restrições de desigualdade. 6.
Estrutura a termo das taxas de juros. I. Título.

CDD: 519.5 (20^a)



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Estatística
Via Washington Luís, Km 235 - C.P.676 - CGC 45358058/0001-40
FONE: (016) 3351-8292 – Email: ppgest@ufscar.br
13565-905 - SÃO CARLOS-SP - BRASIL

FOLHA DE APROVAÇÃO

Aluno(a) : Marcelo Bertini Brocco

DISSERTAÇÃO DE MESTRADO DEFENDIDA E APROVADA EM 21/03/2013
PELA COMISSÃO JULGADORA:

Presidente

Prof. Dr. Márcio Alves Diniz (DEs-UFSCar/Orientador)

1º Examinador

Prof. Dr. Carlos Alberto Ribeiro Diniz (DEs-UFSCar)

2º Examinador

Prof. Dr. Marcio Poletti Laurini (FEA-RP/USP)

Agradecimento

Inicialmente a Deus, que sem a sua vontade nada disso seria possível, por sempre me amparar nas dificuldades e dar forças para continuar meu caminho.

Ao professor, orientador e principalmente um amigo nessa caminhada o Professor Doutor Márcio Alves Diniz, que em 2011 aceitou essa tarefa e com sua dedicação, competência, tempo desprendido em idéias, sugestões, discussões e revisões, colaborou de forma direta na elaboração e conclusão desta dissertação e por ter confiado no meu conhecimento e trabalho.

Incluo também o Prof. Dr. Adriano Polpo de Campos pelos conselhos, sugestões durante o curso e em especial na Banca de Qualificação, o Prof. Dr. Márcio Poletti Laurini por sua participação na Banca Examinadora e sugestões para enriquecimento do trabalho, o Prof. Dr. Carlos Alberto Ribeiro Diniz que me ajudou na escolha em cursar o mestrado, as ajudas durante todo o curso e suas sugestões tanto na banca de Qualificação como na Examinadora para enriquecimento do trabalho.

Agradeço em especial a minha mãe e meu pai que sempre depositaram confiança em mim, sempre me apoiando em todos os momentos da vida, incentivando estudar e dando possibilidades de isso acontecer, sempre dispostos a trazer um lanche ou mesmo ler minha dissertação sem entender nada de estatística. Vocês são responsáveis pela pessoa que sou hoje. Essa vitória não seria possível sem a presença de vocês.

A minha irmã, que sempre foi e será um exemplo na minha vida com sua dedicação, força, vontade de crescer, entre outros atributos que só me mostram a sorte que é ser seu irmão. Agradeço seu apoio sempre nos meus estudos, as conversas e discussões para agregar valor no trabalho e em especial na vida.

Agradeço minha namorada (Bianca Pulgrossi Ferreira), apesar de não estar presente desde o princípio comigo, entrou em minha vida em uma época importante

e com todo seu carinho e paciência me tranquilizava nos maus momentos, ajudava sempre na superação de todas as adversidades e soube compreender todos os momentos que passava.

Esse trabalho também é uma conquista dos meus amigos. Aqui fica meu agradecimento aos meus companheiros de mestrado: Breno Silveira de Andrade, Jonatas Silva do Espírito Santo, Paulo Ricardo Simioni e Tiago Ribeiro Pellegrini, pelas horas de estudos, pelas conversas, por sempre me darem forças para continuar. Ao meu amigo Leonardo Artur Alves Bernardi, por toda a ajuda e aos demais amigos que entendiam o motivo de muitas vezes eu não poder estar com eles.

Agradeço a UFSCar pelo ensino gratuito e de qualidade, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida durante o período de estudo e a todas as pessoas que dão suporte a toda essa estrutura.

Resumo

O objeto de estudo deste trabalho é a curva de taxas de juros, uma importante ferramenta utilizada em decisões financeiras, pois desempenha um papel fundamental na implementação e avaliação de políticas monetárias pelos bancos centrais. Assim sendo, indica as expectativas do mercado quanto ao comportamento futuro das taxas de juros, inflação e atividade econômica. A utilização de um bom modelo e uma boa estimação dos parâmetros do mesmo nos permite representar a curva ajustada o mais próximo da curva real, dessa forma, conseguimos encontrar resultados mais precisos e confiáveis. Neste trabalho estudamos o modelo utilizado para construção das curvas de taxas de juros desenvolvido em 1987 por Nelson e Siegel (1987) e métodos, considerando as restrições do modelo, para obtermos as estimativas dos parâmetros (Mínimos Quadrados Ordinários, Máxima Verossimilhança e Inferência Bayesiana) na versão estática. O modelo de Nelson e Siegel apresenta grande aplicação tanto no Brasil quanto no restante do mundo, pois ele apresenta como características seu caráter parcimonioso nos parâmetros, sua fácil implementação e ajuste eficiente nos diversos formatos que a curva de taxas de juros pode assumir. Por considerarmos as restrições do modelo, encontramos estimativas para os parâmetros do modelo mais seguras e além disso, como principal contribuição deste trabalho, temos uma forma de estimação do parâmetro de tempo (τ) conjuntamente com os demais parâmetros do modelo, sem considerar apenas um valor fixo para ele.

Palavras-chave: *Modelo de Nelson e Siegel. Inferência Bayesiana. Restrições de Desigualdade. Curva de Taxas de Juros. Estrutura a Termo das Taxas de Juros. Estimação de Parâmetros.*

Abstract

The present paper studies the yield curve, an important tool for financial decisions, due to its fundamental role in the implementation and evaluation of monetary policies by the central banks. It also shows market perspectives in relation to the future development of interest rates, inflation and economical activities. Using an adequate model and a reasoned assessment of its parameters enables us to adjust the curve as far as possible to the real curve and hence obtain most precise and trustful results. These results were acquired by studying a model which was developed in 1987 by Nelson and Siegel and used to draw up the yield curve. Considering the model's limitations, different methods were used to attain the estimated parameters, such as Ordinary Least Squares, Maximum Likelihood and Bayesian Inference in the static version. The Nelson-Siegel model is widely used in Brazil and in the rest of the world, due to its economical idea, easy implementation and efficient adjustment into different formats that the yield curve is able to deal with. By considering the restrictions of the model, we found estimations for the parameters of the model safer than other and besides, the main point of this work is an estimation form of parameters of time (τ) together with others parameters of the model without considering one fixed value for it.

Keywords: *Nelson and Siegel Model. Bayesian Inference. Inequality Restrictions. Yield Curve. Term Structure of Interest Rates. Parameter Estimation.*

Sumário

1	Introdução	1
2	Taxa de Juros e Curvas de Taxas de Juros	5
2.1	Taxa de Juros	5
2.1.1	Relações entre taxas	11
2.2	Curva de Taxas de Juros	13
2.2.1	Teorias sobre os formatos das curvas de taxas de juros	14
2.2.2	Formas das Curvas de Taxas de Juros	18
2.2.3	Fatores Responsáveis pela Movimentação da ETTJ	21
3	Modelagem da curva de taxas de juros	23
3.1	Tipos de Modelos	24
3.1.1	Modelos de Regressão	25
3.1.2	Modelos Empíricos	26
3.1.3	Modelos de Equilíbrio	28
3.2	Modelo de Nelson e Siegel	29
3.2.1	Interpretação dos Parâmetros	34
3.3	Esboço da ETTJ pelas estimativas dos parâmetros	39
4	Métodos de Estimação	41
4.1	Mínimos Quadrados Ordinários	45

4.2	Máxima Verossimilhança	49
4.3	Inferência Bayesiana - τ fixo	53
4.3.1	Estudo de Simulação	58
4.4	Inferência Bayesiana - τ estimado	62
4.4.1	Estudo de Simulação	64
5	Aplicação em Dados Reais	66
5.1	Conjunto de Dados	66
5.2	Estimação	70
5.2.1	Mínimos Quadrados Ordinários	70
5.2.2	Máxima Verossimilhança	72
5.2.3	Inferência Bayesiana Aplicada à ETTJ	73
6	Conclusão	88
A	Distribuições Utilizadas	91
	Referências Bibliográficas	93

Capítulo 1

Introdução

No mercado financeiro, certas variáveis contêm informações úteis sobre as condições futuras da economia. Uma dessas variáveis é a taxa de juros, da qual obtemos as informações necessárias para a construção da curva de taxas de juros, também conhecida como estrutura a termo das taxas de juros (ETTJ).

Podemos dizer que a ETTJ é uma importante ferramenta utilizada em decisões financeiras. Em particular, ela é empregada para precificar títulos financeiros, gerenciar riscos e construir portfólios. Além disso, desempenha um papel fundamental na implementação e avaliação de políticas monetárias pelos bancos centrais.

A ETTJ é importante para a condução da política monetária, porque é um indicador das expectativas de mercado sobre a postura adotada pelos bancos centrais, ou seja, ela é capaz de indicar as expectativas dos demais agentes de mercado quanto ao comportamento futuro das taxas de juros, inflação e atividade econômica. Ela contém informações relevantes tanto para os participantes do mercado como para o próprio banco central.

O membro do Conselho do FED - *Federal Reserve*, Banco Central dos Esta-

dos Unidos, Kevin M. Warsh¹, observou que: “as informações contidas nos preços dos ativos, principalmente nas curvas de taxas de juros de mercado, são claramente muito importantes para a boa condução monetária, pois contêm elementos preciosos do ponto de vista das expectativas dos agentes e, portanto, devem ser levadas em consideração pela autoridade quando da tomada de suas decisões. A boa comunicação dos bancos centrais com o mercado possibilita ainda uma melhor interpretação dos sinais emitidos pelos preços dos ativos financeiros”.

Da mesma forma que os bancos centrais, os agentes de mercado tomam decisões com base em informações contidas nas curvas de taxas de juros. Por esse fato, os agentes econômicos devem realizar uma análise precisa do comportamento das curvas de taxa de juros, para assim poderem se antecipar ao que pode ocorrer, seja para maximizar seu lucro, para evitar altas indesejadas da inflação, para não investirem em determinado título ou então para lançarem títulos no mercado.

Notamos que a curva de taxas de juros apresenta grande importância tanto para os bancos centrais como para os demais agentes de mercado. O banco central é responsável por definir as taxas de juros básicas da economia em um certo período de tempo. Com isso ele define, implicitamente, o comportamento da curva de taxas de juros.

Para os diferentes tipos de agentes de mercado, os objetivos na utilização das curvas de taxas de juros são diferentes. Por exemplo: uma empresa que quer lançar um título no mercado utiliza as curvas de taxas de juros para verificar como outras empresas, do mesmo ramo dela, apresentam seus títulos no mercado e assim, ela consegue oferecer títulos com rendimentos atrativos para os investidores. Além disso, para os investidores, as curvas de taxas de juros servem como instrumento para comparar ofertas de títulos de empresas do mesmo setor e verificar qual apre-

¹Warsh, K. M. *Financial Markets and the Federal Reserve*, 21/11/2006, Discurso proferido na Bolsa de Valores de Nova Iorque, disponível em www.federalreserve.gov/newsevents/speech/warsh20061121a.htm

senta o melhor retorno para investir seu capital. Todas essas utilizações dependem de uma modelagem precisa da curva de taxas de juros.

Diversos modelos para as curvas de taxas de juros foram desenvolvidos nas últimas décadas, como exemplo, podemos citar o modelo de Dobbie-Wilkie (Dobbie e Wilkie, 1978 e 1979), o modelo de Carriere (Carriere, 1998), o modelo de Vasicek (Vasicek, 1977), o modelo de Cox, Ingersoll e Ross (Cox, Ingersoll e Ross, 1985), o modelo de Nelson e Siegel (Nelson e Siegel, 1987) e sua versão estendida proposta por Svensson (Svensson, 1994).

Esses modelos são baseados em fórmulas paramétricas e deseja-se que sejam modelos de fácil implementação e parcimoniosos nos parâmetros, ou seja, um modelo que com poucos parâmetros se ajusta com eficiência aos diversos formatos que as curvas podem assumir. Podemos citar como exemplo e que será objeto do nosso estudo, o modelo desenvolvido por Nelson e Siegel, o qual apresenta tais características e é um dos modelos mais utilizados no Brasil e no mundo, utilizado por nove dos treze bancos centrais que fornecem suas estimativas da estrutura a termo para o *Bank of International Settlements* (BIS, 2005).

As diferentes formas funcionais para a curva de taxas de juros são amplamente divulgados na literatura. Contudo, a metodologia empregada na estimação dos parâmetros não é muito discutida. Por exemplo, Ioannides (2003) avaliou diversos modelos que descrevem a ETTJ utilizando o algoritmo BFGS para sua estimação, que representa uma aproximação do método de Newton, enquanto que Csaibok (1998) sugere a utilização do algoritmo de Gauss-Newton para tal procedimento. Manousopoulos e Michalopoulos (2009) compararam diversos métodos tradicionais de otimização não-linear para estimar os parâmetros da curva de taxas de juros. Além de outros métodos de estimação que são citados na literatura e utilizados no mercado.

Neste trabalho estudamos o modelo de Nelson e Siegel em sua versão estática,

ou seja, o ajuste é feito para uma única curva e não para diversas curvas em uma sequência de tempo (modelo na versão dinâmica) e alguns métodos de estimação dos parâmetros do modelo: Mínimos Quadrados Ordinários, Máxima Verossimilhança e a aplicação da Inferência Bayesiana.

O estudo será realizado em seis capítulos, sendo este o primeiro. No capítulo dois, introduzimos os conceitos de taxas de juros e de curvas de taxas de juros. No capítulo três, fazemos a exposição do modelo de Nelson-Siegel. No capítulo quatro discutimos os métodos de estimação pontual (Mínimos Quadrados Ordinários, Máxima Verossimilhança e a Inferência Bayesiana). No capítulo cinco aplicamos os métodos de estimação em um conjunto de dados reais e por fim, no capítulo seis, apresentamos a conclusão do trabalho.

Capítulo 2

Taxa de Juros e Curvas de Taxas de Juros

O modelo proposto por Nelson e Siegel (Nelson e Siegel, 1987) é utilizado na modelagem da curva de taxas de juros para um título específico. Para compreensão do modelo faremos inicialmente, uma breve ilustração sobre taxa de juros (Seção 2.1) e serão apresentadas algumas relações entre taxas. Na sequência é apresentada uma ilustração sobre as curvas de taxas de juros (Seção 2.2). Ainda nesta seção temos os formatos existentes das curvas, teorias que tentam explicar estes formatos e os fatores responsáveis pelo movimento das mesmas. Maiores detalhes sobre essas taxas podem ser encontrados em Fabozzi e Focardi (2004) e Stander (2005).

2.1 Taxa de Juros

Por definição, taxa de juros é a remuneração de um capital aplicado a uma certa taxa durante um determinado período, ou seja, é o capital pago pelo direito de uso de um certo montante durante um intervalo de tempo. São fixados através de uma taxa percentual que sempre se refere a uma unidade de tempo.

Podemos, pela definição, entender juros como um aluguel sobre o dinheiro, ou seja, a taxa cobrada se deve ao fato do tomador do empréstimo ter o direito de usar o dinheiro até o dia do pagamento, enquanto o credor recebe uma compensação por não poder utilizar o montante até essa data.

A diferença entre o montante inicialmente investido e o montante recebido no final do período de investimento é o que chamamos de juros. Normalmente essa diferença é expressa como um percentual sobre o valor emprestado, denominada taxa de juros.

Alguns fatores tais como tempo do empréstimo, risco de inadimplência, inflação, entre outros devem ser considerados para determinar quanto deve ser essa taxa ao alugar o dinheiro.

A taxa de juros pode ser capitalizada de três formas: simples, composta e contínua.

1. Regime de Capitalização Simples:

O juro é calculado unicamente sobre o capital inicial e proporcional ao prazo, ou seja, os juros são pagos no final do período de forma única.

Se temos, portanto, um regime de capitalização simples, o valor futuro resultante da aplicação de um montante inicial P a uma taxa $i\%$ ao período, durante m períodos, é dado por:

$$VF(m) = P(1 + mi).$$

2. Regime de Capitalização Composta:

Nesse regime cobram-se juros sobre juros, ou seja, os juros dos períodos anteriores são acumulados para o período seguinte em forma de progressão geométrica.

Se temos então a aplicação de um montante inicial P a uma taxa $i\%$ ao período, durante m períodos, o valor futuro será dado por:

$$\begin{aligned} VF(1) &= P(1+i) \\ VF(2) &= VF(1)(1+i) = P(1+i)(1+i) = P(1+i)^2 \\ &\vdots \\ VF(m) &= P(1+i)^m. \end{aligned}$$

3. Regime de Capitalização Contínua:

O juro sobre determinado montante aplicado é tratado da mesma forma que em um regime de capitalização composta. A diferença é dada pela frequência de capitalização. No regime de capitalização composta, os juros são capitalizados em períodos discretos (1 dia, 1 mês ou 1 ano, por exemplo) e no caso do regime de capitalização contínua, os juros são capitalizados quase em todo instante de tempo, continuamente.

$$VF(m) = \lim_{m \rightarrow \infty} \left(1 + \frac{i}{m}\right)^m = \exp(mi).$$

Para demonstrar o motivo desse regime de capitalização apresentar a fórmula definida anteriormente, vamos supor que um banco ofereça uma taxa de juros anual de 100% e que investimos inicialmente R\$1,00 nesse banco.

Se no final de 1 ano a capitalização for realizada 1 vez, então:

$$VF(1) = 1(1+1)^1 = R\$2,00.$$

Agora, no caso da capitalização ser realizada 2 vezes no período, temos juros

sobre juros, assim:

$$VF(1) = 1(1 + 1/2)^2 = R\$2, 25.$$

No caso de ser realizada 4 capitalizações no período, temos:

$$VF(1) = 1(1 + 1/4)^4 = R\$2, 44.$$

E se todo mês fosse realizada uma capitalização, então:

$$VF(1) = 1(1 + 1/12)^{12} = R\$2, 61.$$

Isso nos leva a imaginar que se a capitalização fosse realizada a cada instante de tempo, ao final do período, esperamos ter um valor muito maior em relação ao valor investido inicialmente. Contudo, compor o juro continuamente em todo o instante de tempo nos leva a uma sucessão que converge para a base do logaritmo natural ($e \cong 2,718$), ou seja, o máximo recebido para um montante inicial de R\$1,00 seria de aproximadamente R\$2,72 em um ano, mesmo com capitalizações infinitesimais.

Por esse fato, podemos dizer que para um regime de capitalização contínua basta utilizarmos a fórmula anteriormente definida.

Neste trabalho utilizamos também os seguintes conceitos de taxas de juros: *taxas forward* (taxas futuras), *spot rate* (taxas à vista) e *fator de desconto*.

1. *Taxas Forwards*:

Definimos *taxa forward* como sendo a taxa de juros obtida a partir da *taxa à vista* para períodos de tempo no futuro, ou seja, é permitido fixar no futuro

uma taxa de juros definida no presente. Denotaremos *taxas forward* por $f_t(m)$, onde m denota a maturidade.

Podemos dizer que o comprador compromete-se a pagar ao vendedor, em uma data futura uma taxa de juros fixada no presente, recebendo em troca uma taxa de juros variável. Assim, se a taxa de referência usada subir entre a data em que o contrato é válido e a data de fixação da taxa, o comprador terá dinheiro a receber na data de fixação. Se as taxas baixarem, o comprador terá dinheiro a pagar na data de fixação. É garantida, assim, sempre a mesma taxa efetiva de financiamento nessa data.

Para exemplificar a utilização da *taxa forward*, utilizaremos um caso em que a *taxa à vista* para um investimento de um ano é de 10% a.a. e para dois anos é de 11% a.a.. Ao final de um ano com investimento de 10% a.a., o investidor deseja investir novamente no segundo ano a uma determinada taxa que apresente retorno equivalente se tivesse investido a uma taxa de 11% a.a. para um período de dois anos, para isso a *taxa forward* que deve ser contratada para este investimento, supondo valor inicial de R\$1,00, é de:

$$\begin{aligned} \exp(0, 10) \exp(i) &= \exp(0, 11 \times 2) \\ \Rightarrow 0, 10 + i &= 0, 22 \Rightarrow i = 0, 22 - 0, 10 \Rightarrow i = 0, 12. \end{aligned}$$

Desta forma, para que o investidor obtenha o montante desejado, a *taxa forward* a ser contratada para o segundo ano é de 12% a.a..

2. Spot Rate:

Definimos *taxa à vista* ($r_t(m)$) como sendo a taxa exigida pelo mercado

no momento da compra do título até seu vencimento contraído também no presente, ou seja, na mesma data.

3. *Fator de Desconto:*

Fator de desconto é o fator com que descontamos certo montante futuro a fim de torná-lo comparável em valores com o dinheiro presente. Esse fator assume valores no intervalo $[0,1]$. Denotaremos o fator de desconto por $f_d(m)$, onde m denota a maturidade.

Podemos construir a curva de fator ou função de desconto utilizando um conjunto de fatores de desconto contra os vencimentos dos títulos. Normalmente a forma de representação é uma função exponencial negativa, na qual temos que $f_d(0)=1$ e os valores para os próximos vencimentos decaem para próximo de zero. Esta representação é observada pela Figura (2.1).

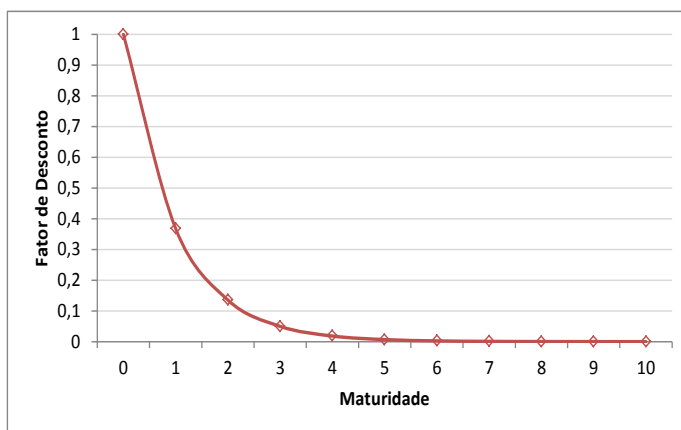


Figura 2.1: Fator de desconto com taxa $x\%$ ao período.

Considerando os regimes de capitalização simples, composto e contínuo, definidos anteriormente, pode-se calcular o fator de desconto através das seguintes relações:

$$f_d(m) = \begin{cases} \frac{1}{(1+i)^m}, & \text{no caso de um regime de capitalização composta.} \\ \frac{1}{(1+im)}, & \text{no caso de um regime de capitalização simples.} \\ \exp(-im), & \text{no caso de um regime de capitalização contínua.} \end{cases}$$

2.1.1 Relações entre taxas

Existem relações entre as diferentes expressões das taxas de juros, possibilitando ao investidor trabalhar com a forma que mais lhe agrada.

Para isso, se tivermos uma *taxa à vista*, podemos obter o *fator de desconto*, por meio da seguinte relação:

$$f_d(m) = \exp(-mr_t(m)). \quad (2.1)$$

Como podemos observar pela equação (2.1), temos aqui o *fator de desconto* definido com relação ao regime de capitalização contínuo, como vimos na seção anterior, onde a taxa de juros i é representado pela *taxa à vista* ($r_t(m)$).

Além disso, podemos a partir do *fator de desconto* calcular a *taxa forward*.

$$f_t(m) = \frac{-f'_d(m)}{f_d(m)}. \quad (2.2)$$

Abaixo exemplificamos essas relações.

Inicialmente obtemos a taxa à vista através da função $r_t(m) = 2m$.

Tabela 2.1: Determinando as taxas à vista

Período	0,2	0,4	0,6	0,8
$r_t(m) = 2m$	0,4	0,8	1,2	1,6

Aplicando a equação (2.1), transformamos as *taxas à vista* em *fator de desconto*:

Tabela 2.2: Determinando o fator de desconto em relação as taxas à vista

Período	0,2	0,4	0,6	0,8
$f_d(m) = \exp(-2m^2)$	0,92	0,73	0,49	0,28

Por fim, para obtermos a *taxa forward* aplicamos a equação (2.2). Temos que $-f'_d(m) = 4m \exp(-2m^2)$ e assim $f_t(m) = 4m$. Desta forma obtemos as seguintes taxas:

Tabela 2.3: Determinando as taxas forward em relação ao fator de desconto

Período	0,2	0,4	0,6	0,8
$f_t(m) = 4m$	0,8	1,6	2,4	3,2

Com isso concluímos que as transformações não são difíceis de serem realizadas, desta forma, o investidor consegue trabalhar com a forma que considere mais fácil o manuseio e entendimento. Já as conclusões que são encontradas pelos diferentes tipos de definição da taxa de juros podem ser diferentes, pois se considerarmos a interpolação, por exemplo, realizada antes do início da conversão das taxas as conclusões serão as mesmas, por outro lado se essa interpolação, por exemplo, for realizada após a conversão da taxa as conclusões podem ser diferentes.

2.2 Curva de Taxas de Juros

Com as informações sobre as taxas de juros, podemos obter uma representação gráfica da ETTJ. Essa curva revela a dependência da taxa de juros em relação aos prazos de vencimentos de um título, isto é, ela informa para cada prazo de vencimento quanto o mercado demanda de juros, revelando assim se as taxas de curto prazo são maiores que as de longo prazo.

A construção da curva de taxas de juros é realizada para títulos específicos de características idênticas, diferenciados apenas pelo prazo de seus vencimentos. Com a utilização de apenas um tipo de título conseguimos isolar o fator da estrutura a prazo dos demais fatores determinantes da variabilidade das taxas de juros (risco, tributação, condições de pagamentos, etc.). Através dessa curva podemos adotar uma tática de compra e venda de um determinado título segundo algum critério estipulado.

Para exemplificar a representação gráfica da ETTJ utilizaremos o seguinte conjunto de dados da Tabela (2.4).

Tabela 2.4: Dados para exemplificação da ETTJ

Período (ano)	Taxa de juros (% ao ano)	Período (ano)	Taxa de juros (% ao ano)
1	5,88	11	5,83
2	6,25	12	5,8
3	6,16	13	5,77
4	6,07	14	5,75
5	6,01	15	5,72
6	5,97	16	5,69
7	5,94	17	5,67
8	5,91	18	5,64
9	5,89	19	5,62
10	5,86	20	5,59

Com o conjunto de dados podemos construir uma representação gráfica da curva de taxas de juros.

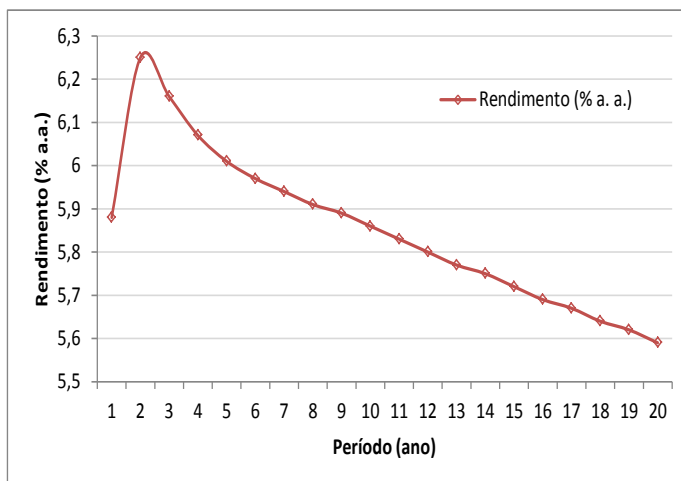


Figura 2.2: Exemplo de representação da ETTJ

A Figura (2.2) mostra a representação da ETTJ. No eixo vertical temos os valores das taxas de juros, que variam entre 5,5% e 6,3% ao ano. No eixo horizontal temos as maturidades para o título variando de 1 à 20 anos. Notamos que a curva é crescente para os dois primeiros vencimentos e a partir desta maturidade a curva passa a ser decrescente, atingindo seu menor valor no vigésimo vencimento.

2.2.1 Teorias sobre os formatos das curvas de taxas de juros

A Figura (2.2) nos apresentou um determinado formato da curva de taxas de juros. Porém, quando atuamos no mercado financeiro podemos verificar alguns formatos específicos, definidos como normal, invertido, plano e ondulado.

Pelo fato das curvas de taxas de juros assumirem quase certamente um desses quatro formatos, algumas teorias surgiram para tentar explicá-los. Em geral, podemos destacar três teorias que apresentam grande aceitação e aplicação no mercado financeiro: Teoria da Segmentação do Mercado, Teoria da Pura Expectativa e Teoria do Prêmio por Liquidez.

Cada uma dessas teorias, além de ilustrar as diferentes formas que uma curva pode assumir, define os três fatores empíricos mais importantes para sua classificação, que são:

1. As taxas de juros com diferentes vencimentos movimentam-se juntas no tempo.
2. Quando as taxas de juros de curto prazo estão baixas, é mais provável que as curvas de taxas de juros sejam normais. Já, quando as taxas de juros de curto prazo estão altas é mais provável que as curvas de taxas de juros assumam a forma invertida.
3. As curvas de taxas de juros geralmente são crescentes, ou seja, apresenta o formato classificado como normal.

A seguir apresentamos, de forma resumida, cada uma das teorias citadas.

a) Teoria da Pura Expectativa

Também conhecida como Teoria das Expectativas Não Enviesadas, formulada por Irving Fisher ¹, essa teoria nos mostra que a taxa de juros de longo prazo se constitui numa média geométrica das taxas de juros de curto prazo correntes.

$$(1 + i_{m+1})^m = (1 + i_1)(1 + i_2) \dots (1 + i_m), \quad (2.3)$$

¹*Appreciation and Interest*, The American Economic Association, Vol.XI, No.4, 8/1896, pp. 331:442

sendo i_1 a taxa de juros do 1º período, ... e i_m a taxa de juros do período m .

Para ilustrar essa Teoria, utilizaremos o seguinte exemplo:

Considere que os ativos com maturidade de três anos são negociados no mercado de acordo com as seguintes informações: Um título com vencimento em um ano rende uma taxa de juros de 10% ao ano, um título com vencimento em dois anos rende 12% a cada ano e por fim um título com vencimento em três anos rende uma taxa de juros de 15% a cada ano.

A estrutura temporal das taxas de juros para os próximos anos é encontrada conforme os seguintes cálculos:

1º Ano: 10% ao ano.

Dado que no primeiro ano tivemos uma taxa de juros de 10% a.a., para obtermos um rendimento no 2º ano de mesmo valor que quando temos um título com vencimento em dois anos e taxa de juros de 12% a.a. devemos contratar uma taxa para o segundo ano de:

2º Ano:

$$\begin{aligned}(1, 12)^2 &= 1,2544 \\ 1,2544 &= [(1, 10)(1 + i)] \rightarrow i = [(1,2544/1,10) - 1] \times 100 \rightarrow \\ i &= 14,04\% \text{ ao ano.}\end{aligned}$$

Dado que no primeiro ano tivemos uma taxa de juros de 10% a.a. e no segundo ano uma taxa de juros de 14,04% a.a., para obtermos um rendimento no 3º ano de mesmo valor que quando temos um título com vencimento em três anos e taxa de juros de 15% a.a., devemos contratar uma taxa para o terceiro ano de:

3º Ano:

$$(1,15)^3 = 1,5208$$

$$1,5208 = [(1,10)(1,1404)(1+i)] \rightarrow i = [1,5208/(1,10 \times 1,1404) - 1] \times 100$$

$$i = 21,24\% \text{ ao ano.}$$

Ou seja, se desejamos trabalhar com títulos com vencimentos anuais e obtermos o mesmo rendimento recebido com títulos com vencimento em dois e três anos, temos que, a cada ano, contratar uma taxa anual maior do que a taxa contratada para os títulos com vencimentos mais longos.

b) Teoria do Prêmio por Liquidez

Proposta por John R. Hicks ², é baseada no pressuposto que os investidores têm preferência por títulos de curto prazo, assim como a Teoria da Segmentação do Mercado.

Essa teoria considera que títulos de curto prazo estão sujeitos a riscos menores que títulos de longo prazo. É razoável que o risco aumenta à medida que aumenta o prazo de vencimento, dado que o prazo e a volatilidade nos preços são diretamente relacionados.

Para incentivar os investidores a optarem por títulos de longo prazo é oferecido um prêmio pela liquidez. Assim, as taxas a longo prazo são dadas pelas taxas futuras esperadas de curto prazo mais um prêmio. O que os tornam mais atrativos aos investidores.

$$(1 + i_{m+1})^m = (1 + i_1)(1 + i_2) \dots (1 + i_m) + \text{Prêmio} \quad (2.4)$$

² *Value and Capital an inquiry into some fundamental principles of economic theory*, Clarendon Press in Oxford, 1946

c) Teoria da Segmentação do Mercado

Sugerida por J. M. Culbertson ³, refere-se ao fato dos investidores terem sua preferência definida por títulos de curto prazo ou de longo prazo. De forma que a oferta e a demanda são os fatores que determinam o valor da taxa de juros.

Assim notamos que existe uma segmentação do mercado em dois grupos de investidores: os que preferem títulos de curto prazo e os que preferem títulos de longo prazo e cada um desses grupos tem seus próprios compradores e vendedores.

Em geral, os investidores apresentam preferência por títulos de curto prazo de vencimento, pois dessa forma eles obtêm uma maior liberdade em investimentos futuros. Assim, a demanda por títulos de curto prazo é maior que a por títulos de longo prazo e por isso as taxas de juros de curto prazo são menores.

As três teorias são aceitas e utilizadas para explicação dos formatos das curvas de taxas de juros. Contudo, cada uma delas explica fatos empíricos que as outras não podem explicar.

2.2.2 Formas das Curvas de Taxas de Juros

Vimos que, em geral, as curvas de taxas de juros apresentam um dos seguintes formatos: normal, invertido, ondulado e plano. Além disso, essas curvas apresentam características específicas, como:

a) Curva Normal ou Positiva

³*The Term Structure of Interest Rates*, The Quarterly Journal of Economics, Vol.71, No.4, 11/1957, pp. 485:517

Apresenta as taxas de juros para títulos de longo prazo de vencimento com rendimentos superiores quando comparados com os títulos de vencimentos de curto prazo. A explicação deste tipo de curva pode ser atribuído à fatores como por exemplo o risco e o tempo do empréstimo.

Uma ilustração do formato da curva normal pode ser observado na Figura (2.3), onde no eixo horizontal temos as maturidades, ou prazo de vencimento do título, e no eixo vertical temos as taxas de juros (rendimento), ou seja, quanto rende o título de determinada maturidade.

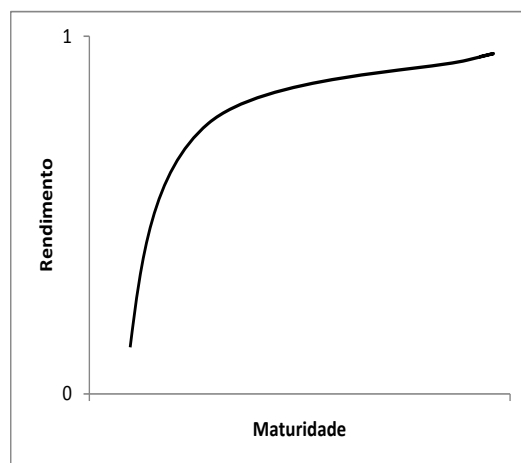


Figura 2.3: Curva Normal

b) Curva Invertida ou Negativa

Este tipo de curva apresenta a ideia contrária em relação à curva normal (Figura (2.4)), ou seja, neste tipo de curva os títulos de curto prazo apresentam rendimentos superiores quando comparados aos de longo prazo.

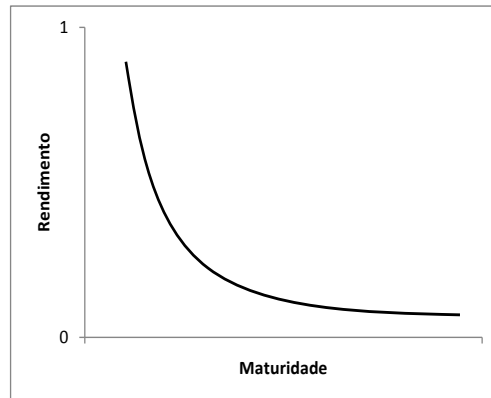


Figura 2.4: Curva Invertida

c) Curva Ondulada

Tem por característica apresentar títulos de curto e longo prazo com rendimentos semelhantes, enquanto que os títulos de médio prazo de vencimento apresentam rendimentos superiores ou inferiores (Figura (2.5)). Este tipo de curva em geral, é observado quando a economia passa por um período de uma recessão ou antes de uma expansão econômica.

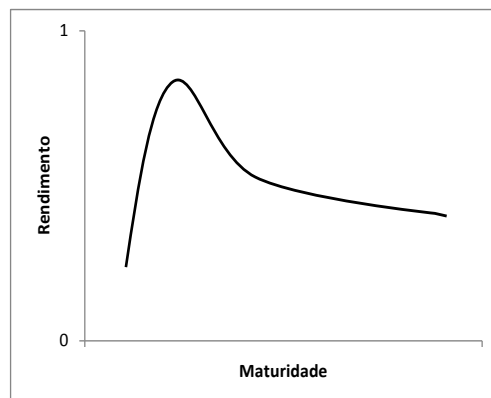


Figura 2.5: Curva Ondulada

d) Curva Plana

Ao contrário de todas as curvas vistas anteriormente, este tipo de curva apresenta como característica todos os prazos de vencimento com rendimentos iguais, independente de termos um título de curto, médio ou longo prazo de vencimento. Em geral, esta curva aparece quando há sinal de incerteza no mercado sobre o futuro da economia. Assim, ela pode em um curto período, se transformar em um dos outros três tipos de curvas vistos anteriormente.

2.2.3 Fatores Responsáveis pela Movimentação da ETTJ

Conforme os preços dos títulos as suas taxas de juros implícitas variam e as curvas se modificam. Essa dinâmica das curvas pode ser explicada por alguns fatores. Em geral, essa associação dos parâmetros dos modelos aos fatores responsáveis pela movimentação da ETTJ é feita normalmente para os modelos em suas versões dinâmicas, contudo a associação nos casos de versões estáticas não estão incorretas.

Os três principais fatores são:

- **Nível de Juros:**

Refere-se aos movimentos paralelos da curva de taxas de juros (deslocamentos para cima ou para baixo), ou seja, quando todos os rendimentos se movem em uma mesma direção com uma mesma magnitude.

- **Inclinação:**

A inclinação pode ser positiva ou negativa em função de sua maturidade, ou seja, é a primeira derivada dos rendimentos em relação aos prazos. É o fator que determina para qual sentido a curva irá, isto é, se o fator nível de juros refere-se a um movimento paralelo para cima ou para baixo, fazendo com que a curva rotacione no sentido horário ou antihorário.

- **Curvatura:**

É a segunda derivada dos rendimentos em relação aos prazos e é o que define a velocidade na qual as taxas crescem (no caso da inclinação positiva), ou decrescem (no caso da inclinação negativa). Ela faz com que taxas de prazos intermediários subam enquanto que, tanto as taxas curtas como as longas, caiam ou vice-versa. Esse movimento é responsável por aumentar ou diminuir a concavidade da curva.

Segundo um estudo realizado por Litterman e Scheinkman (1991) com títulos do tesouro americano, conclui-se que os três fatores vistos anteriormente explicam cerca de 97% da movimentação da curva de taxas de juros sendo que, o fator nível de juros explica 79%, o fator inclinação cerca de 13% e por fim o fator curvatura é responsável por 5% da movimentação.

Capítulo 3

Modelagem da curva de taxas de juros

A ETTJ é importante tanto para os agentes de mercado como para o próprio banco central. O banco central é quem define as taxas de juros básicas vigentes para um certo período. Dessa forma, ele pode controlar o formato que as curvas de taxas de juros assumem, ou seja, definir qual dos formatos apresentados na Seção 2.2.2, ela irá assumir.

Para os demais agentes de mercado, o objetivo de utilizar as curvas de taxas de juros são diferentes. Por exemplo, uma empresa que deseja lançar no mercado seus títulos, consegue com base na comparação das curvas de outras empresas do mesmo segmento, saber se aquele momento é ou não favorável para o lançamento de seus títulos e além disso, definir um preço atraente aos investidores. Outra utilização acontece com os investidores que, podem comparar entre as diversas ofertas de títulos qual o mais rentável para o tipo de investimento que deseja realizar.

Para ilustrar essa ideia, considere um investidor que está interessado em títulos da empresa A ou da empresa B. Sabemos que esse investidor tem preferência por títulos de curto prazo de vencimento. Ele deseja saber qual oferta é mais rentável:

títulos da empresa A de 2 anos ou da empresa B de 3 anos. Para tomar sua decisão ele ajusta as curvas de taxas de juros para estas duas empresas e compara os vencimentos nas mesmas maturidades.

Pela necessidade de se obter resultados precisos aplica-se a modelagem das curvas de taxas de juros. Esta modelagem deve apresentar os valores ajustados o mais próximo possível dos valores reais observados. Um mau ajuste dessa curva pode resultar em conclusões errôneas e levar os agentes de mercado à prejuízos ou a retorno abaixo do esperado.

Neste capítulo apresentaremos alguns modelos sugeridos na literatura para modelarmos a ETTJ (Seção 3.1). Na Seção 3.2 será apresentado mais detalhadamente o modelo de Nelson e Siegel, objeto de estudo deste trabalho e a interpretação dos parâmetros do modelo (Parâmetros de Regressão e Parâmetro de Tempo). Por fim, a Seção 3.3 mostrará um esboço da curva pelas estimativas dos parâmetros.

3.1 Tipos de Modelos

A literatura sugere alguns modelos para a curva de taxas de juros divididos em algumas classes. Essas classes, dependendo da referência, podem ser divididas em duas ou três classes de acordo com suas características.

Em geral, quando temos três classes de modelos, denominamos elas como: Modelos de Regressão, Modelos Empíricos e Modelos de Equilíbrio. E no caso de duas classes, temos que os Modelos de Regressão e Modelos Empíricos são denominados Modelos Estatísticos ou Modelos de Não Arbitragem e a outra classe continua sendo denominada Modelos de Equilíbrio ou Modelos de Arbitragem.

Os modelos mencionados nas sub-seções posteriores e outros exemplos de modelos pertencentes a cada classe, são encontrados em Stander (2005, pg. 33:52).

3.1.1 Modelos de Regressão

Um modelo de regressão consiste no uso de uma amostra de dados econômicos para obter informações sobre relações entre variáveis. Esses modelos de regressão consideram um gráfico de dispersão dos rendimentos contra o prazo de vencimento de um título específico. Dessa forma, apenas o vencimento do título e o seu rendimento, são utilizados para modelar a curva de taxa de juros.

O critério de seleção para escolha do melhor ajuste, geralmente é a Soma de Quadrados dos Resíduos (SQR).

$$SQR = \sum_{i=1}^n (R_i - \hat{R}_i)^2, \quad (3.1)$$

em que R_i se refere aos rendimentos da taxa de juros definido pelo mercado e \hat{R}_i aos rendimentos da taxa de juros ajustada pelo modelo para a i -ésima maturidade.

A SQR é uma medida de distância entre os valores observados e os valores estimados do rendimento. Por se tratar de uma diferença ao quadrado, a medida só apresenta resultados positivos e, com isso, o melhor ajuste será dado pela curva que apresentar o menor valor da SQR. Essa será a curva que mais se aproxima, segundo esse critério, dos valores observados.

Dentro dessa classe de modelos podemos citar os modelos:

- Modelo de Bradley-Crane

$$\ln(1 + R_i) = \beta_0 + \beta_1 m_i + \beta_2 \ln(m_i),$$

em que R_i é o retorno no vencimento i , m_i é a maturidade em anos no vencimento i e β_0 , β_1 e β_2 são os parâmetros de regressão.

- Modelo de Dobbie-Wilkie

$$R_i = \beta_0 + \beta_1 \exp(-\alpha_1 m_i) + \beta_2 \exp(-\alpha_2 m_i),$$

em que R_i é o retorno no vencimento i , m_i é a maturidade em anos no vencimento i , β_0 , β_1 e β_2 são os parâmetros de regressão linear e α_1 e α_2 são os parâmetros da regressão não-linear.

3.1.2 Modelos Empíricos

A classe de modelos empíricos¹ foi desenvolvida para superar algumas deficiências que os modelos de regressão apresentavam. Uma das principais vantagens dessa classe em relação aos Modelos de Regressão é o fato de apresentar maior flexibilidade na modelagem da curva de taxa de juros, permitindo um ajuste mais preciso da curva. Podemos tornar o ajuste da curva cada vez mais preciso à medida que aumentamos o grau do polinômio envolvido na modelagem da curva. Contudo, à medida que aumentamos o grau do polinômio do modelo aumentamos a dificuldade para a estimação dos parâmetros e os tornamos cada vez menos parcimoniosos.

O critério de seleção do melhor ajuste para a curva de taxa de juros é o mesmo utilizado em modelos de regressão, a SQR.

Os modelos empíricos são construídos através do *fator de desconto* (Seção (2.1)), ou seja, o preço de um título de renda fixa é igual ao fluxo de caixa futuro prometido pelo emissor trazido a valor presente pelo *fator de desconto*. Através do *fator de desconto*, podemos encontrar a função base definida por k funções linearmente independentes que pode ser escrita como:

¹Não confundir esses modelos com os modelos estatísticos baseados na função de distribuição empírica

$$f_d(m) = \beta_0 + \sum_{j=1}^k \beta_j f_j(m), \quad (3.2)$$

em que os β_j são os parâmetros a serem estimados, $f_j(m)$ é a função base e m a maturidade.

A função base e o valor de k são importantes para uma boa qualidade do ajuste do modelo.

McCulloch (1971, 1975) foi o pioneiro no uso de funções matemáticas (splines polinomiais), para ajustar empiricamente a função de desconto em determinados momentos do tempo. Na sequência, em 1982, Vasicek e Fong, modelaram a curva de desconto usando splines exponenciais. Depois disso, várias outras funções foram propostas para ajustar a curva de taxa de juros.

Alguns exemplos de modelos pertencentes a essa classe são:

- Modelo de Nelson e Siegel

$$R_i(m) = \beta_0 + \beta_1 \frac{[1 - \exp(-m_i\tau)]}{(m_i\tau)} + \beta_2 \left[\frac{[1 - \exp(-m_i\tau)]}{(m_i\tau)} - \exp(-m_i\tau) \right],$$

em que R_i é o retorno no vencimento i , m_i é a maturidade em anos no vencimento i , β_0 , β_1 e β_2 são os parâmetros de regressão e τ é o parâmetro de tempo.

- Modelo de Svensson

$$R_i(m) = \beta_0 + \beta_1 \frac{[1 - \exp(-m_i\tau_1)]}{(m_i\tau_1)} + \beta_2 \left[\frac{[1 - \exp(-m_i\tau_1)]}{(m_i\tau_1)} - \exp(-m_i\tau_1) \right] + \beta_3 \left[\frac{[1 - \exp(-m_i\tau_2)]}{(m_i\tau_2)} - \exp(-m_i\tau_2) \right],$$

em que R_i é o retorno no vencimento i , m_i é a maturidade em anos no vencimento i , β_0 , β_1 , β_2 e β_3 são os parâmetros de regressão e τ_1 e τ_2 são os

parâmetros de tempo.

3.1.3 Modelos de Equilíbrio

Para construção destes modelos assume-se que as taxas de juros seguem uma equação diferencial estocástica. Assim, determina-se a forma funcional da função de desconto que explica a dinâmica da estrutura da curva de taxa de juros ao longo do tempo. Como estes modelos tem algumas restrições, o ajuste aos dados reais não é perfeito, o que resulta em curvas menos precisas do que as estimadas pelos Modelos Empíricos.

Os modelos de equilíbrio, por definição, utilizam variáveis. Os modelos mais básicos têm uma única variável econômica que explica o nível da curva de rendimentos da taxa de juros. Estas variáveis nesses modelos são denominadas fatores.

Dentro dessa classe de modelos podemos citar:

- Modelo de Vasicek

$$dR = \alpha(\gamma - R)dt + \sigma dZ.$$

em que R é a taxa livre de risco, γ é a taxa de longo prazo esperada, α é a taxa de reversão média, σ é a taxa de volatilidade e dZ é um termo estocástico normalmente distribuído.

- Modelo de Cox, Ingersoll e Ross

$$dR = \alpha(\gamma - R)dt + \sigma r^\beta dZ.$$

segue do modelo de Vasicek, porém é proposto o processo de risco (volatilidade) neutro.

3.2 Modelo de Nelson e Siegel

O modelo de Nelson e Siegel é amplamente utilizado em mercados financeiros pelo mundo, pois ajusta com eficiência as diversas formas que as curvas de taxas de juros podem assumir. Ele é um dos modelos que melhor descreve as taxas de juros do mercado brasileiro, conhecido por apresentar uma ETTJ de difícil modelagem.

O modelo foi desenvolvido por Charles R. Nelson e Andrew F. Siegel em 1987, e é conhecido como modelo de Nelson e Siegel. Ele está classificado na classe de modelos empíricos, possui quatro parâmetros e por isso é considerado parcimonioso, isto é, um modelo que com poucos parâmetros consegue se adequar aos diversos formatos que as curvas de taxas de juros podem assumir.

Esse modelo está classificado dentro da classe dos modelos empíricos e sua definição é dada em termos do *fator de desconto*. Contudo, como visto na Seção (2.1.1), através das relações existentes entre as taxas de juros podemos reescrever o modelo em função da *taxa à vista* e da *taxa forward*.

A origem de sua forma funcional está em equações diferenciais ordinárias de segunda ordem. As soluções gerais destas equações geram rapidamente as formas típicas das curvas de taxas de juros.

Neste caso, se temos uma equação diferencial ordinária de segunda ordem dada por:

$$af_t''(m) + bf_t'(m) + cf_t(m) = d, \quad (3.3)$$

em que a , b , c e d são constantes reais conhecidas e $a \neq 0$.

Para resolvermos a equação (3.3), primeiramente, procuramos soluções da forma $f_t(m) = \exp(\lambda m)$ para a equação homogênea. Ela pode ser obtida substituindo esse resultado na equação (3.3), assim obtêm-se:

$$\begin{aligned}
 a\lambda^2 \exp(\lambda m) + b\lambda \exp(\lambda m) + c \exp(\lambda m) &= 0 \\
 \exp(\lambda m)(a\lambda^2 + b\lambda + c) &= 0.
 \end{aligned} \tag{3.4}$$

Nosso objetivo passa a ser encontrar a solução não-trivial ($\lambda \neq 0$):

$$a\lambda^2 + b\lambda + c = 0. \tag{3.5}$$

Se da equação (3.5) tivermos como solução duas raízes reais e distintas, a solução homogênea para a equação diferencial ordinária de segunda ordem fica sendo:

$$f_t(m) = A_0 + A_1 \exp(\lambda_1 m) + A_2 \exp(\lambda_2 m),$$

em que A_0 , A_1 e A_2 são constantes que serão determinadas através de condições iniciais.

Caso a solução da equação (3.5) seja dada por raízes reais iguais, a solução homogênea da equação será da forma:

$$f_t(m) = A_0 + A_1 \exp(\lambda m) + A_2 m \exp(\lambda m),$$

em que A_0 , A_1 e A_2 são constantes que serão determinadas através de condições iniciais.

Inicialmente Nelson e Siegel propuseram o caso com duas raízes reais e distintas, mas posteriormente foram detectados alguns problemas. Entre eles podemos citar a não parcimoniosidade do modelo e a dificuldade no ajuste das formas típicas da curva de taxa de juros (Seção 2.2.2). Sendo assim, o modelo foi ajustado para o caso de raízes reais iguais, no qual o mesmo passa a ser parcimonioso, gera as formas

típicas da curva de taxa de juros e os resultados apresentados são satisfatórios.

Para definição do modelo foi utilizado então o caso em que a equação apresentava raízes reais e iguais. Além disso, para chegar à forma funcional do modelo foi proposta a resolução do seguinte problema de valor inicial (PVI):

$$\begin{cases} f_t(0) = \beta_0 + \beta_1 \\ f_t'(0) = \beta_1\lambda - \beta_2\lambda \\ f_t''(0) = \beta_1\lambda^2 - 2\beta_2\lambda^2. \end{cases}$$

Para resolvermos o PVI devemos obter as derivadas da função $f_t(m)$ para o caso de raízes reais e iguais.

$$\begin{cases} f_t(m) = A_0 + A_1 \exp(\lambda m) + A_2 m \exp(\lambda m) \\ f_t'(m) = \lambda A_1 \exp(\lambda m) + A_2 \exp(\lambda m) + A_2(m\lambda) \exp(\lambda m) \\ f_t''(m) = \lambda^2 A_1 \exp(\lambda m) + \lambda A_2 \exp(\lambda m) + \lambda A_2 \exp(\lambda m) + A_2(m\lambda^2) \exp(\lambda m) \end{cases}$$

Utilizando o PVI e as equações anteriores calculadas, temos:

$$\begin{cases} f_t(0) = A_0 + A_1 = \beta_0 + \beta_1 \\ f_t'(0) = \lambda A_1 + A_2 = \lambda\beta_1 - \lambda\beta_2 \\ f_t''(0) = \lambda^2 A_1 + 2\lambda A_2 = \beta_1\lambda^2 - 2\beta_2\lambda^2 \end{cases}$$

E podemos verificar facilmente que:

$$\begin{cases} A_0 = \beta_0 \\ A_1 = \beta_1 \\ A_2 = -\beta_2\lambda \end{cases}$$

Com isso temos que o modelo estático de Nelson e Siegel é dada por:

$$f_t(m) = \beta_0 + \beta_1 \exp(m\lambda) - \beta_2[(m\lambda) \exp(m\lambda)].$$

A forma funcional usada na literatura considera a transformação $\lambda = -\tau$. Assim, a forma mais conhecida e utilizada é:

$$f_t(m) = \beta_0 + \beta_1 \exp(-m\tau) + \beta_2[(m\tau) \exp(-m\tau)], \quad (3.6)$$

em que β_0 , β_1 , β_2 e τ são os parâmetros do modelo que devemos estimar e m é a maturidade do título.

Com as seguintes restrições:

$$\begin{cases} \hat{\beta}_0 \geq 0. \\ \hat{\beta}_0 + \hat{\beta}_1 \geq 0. \\ \hat{\tau} \geq 0. \end{cases}$$

Na sequência deste trabalho será explicado o motivo destas restrições serem impostas ao modelo.

Temos que o modelo foi definido em termos das *taxas forwards*, mas podemos definí-lo em relação ao *rendimento até o vencimento* (*yield to maturity*) (YTM), neste trabalho utilizaremos a denominação rendimento para referirmos a YTM.

O *Rendimento até o vencimento* é a taxa de rendimento que o comprador do título recebe se ficar com o título até o vencimento. Alguém que compra um título com o preço de mercado de hoje recebe a taxa implícita de juros desse título se ficar com o título até seu vencimento.

Para exemplificar, suponha um título pré-fixado que não paga juros (cupom) e vence daqui 30 anos. Seu valor de face é de R\$100 e YTM anual de 10% a.a..

O preço do título é dado por $VP = \frac{P}{(1+i)^m} = \frac{100}{(1,1)^{30}} = R\$5,73$ e, se esperarmos os trinta anos, receberemos R\$100. Contudo, não sabemos o que ocorrerá neste

intervalo de tempo. Suponha que após 10 anos do lançamento do título, a taxa caia para 7% a.a.. Se o comprador 1 resolve vender o título a um comprador 2, esse comprará o título por $100/(1,07)^{20} = R\$25,84$ com uma de YTM anual de 7% a.a.. Para o comprador 1, que vendeu o título após 10 anos, o lucro foi de R\$20,11, gerando um retorno de 16,25% a.a. ao invés de 10% a.a. que receberia no final dos 30 anos se permanecesse com o título até o vencimento.

É possível estabelecer a seguinte relação entre a YTM e as *taxas forward*:

$$R(m) = \int_0^m \frac{f_t(x)dx}{m}. \quad (3.7)$$

Utilizaremos a forma do modelo definida em função das taxas forwards (equação (3.6)), e aplicando a equação anterior, encontramos o modelo de Nelson e Siegel em termos do rendimento até o vencimento.

$$\begin{aligned} R(m) &= \int_0^m \frac{f_t(x)dx}{m} = \frac{1}{m} \int_0^m \beta_0 + \beta_1 \exp(-x\tau) + \beta_2(x\tau) \exp(-x\tau) dx \\ &= \frac{1}{m} \int_0^m \beta_0 dx + \frac{1}{m} \int_0^m \beta_1 \exp(-x\tau) dx + \frac{1}{m} \int_0^m \beta_2(x\tau) \exp(-x\tau) dx. \end{aligned} \quad (3.8)$$

Para auxiliar na resolução dessa integral utilizaremos a regra de integração por partes. Detalhes podem ser encontrados em Costa Neto, A. S. (Cálculo II, 2006, pg.16).

Resolvendo a equação anterior, obtemos o modelo definido em termos do *YTM*.

$$\begin{aligned} &= \frac{1}{m} \int_0^m \beta_0 dx + \frac{1}{m} \int_0^m \beta_1 \exp(-x\tau) dx + \frac{1}{m} \int_0^m \beta_2(x\tau) \exp(-x\tau) dx \\ &= \beta_0 + \frac{\beta_1(1 - \exp(-m\tau))}{m\tau} + \beta_2 \left(-\exp(-m\tau) - \frac{\exp(-m\tau)}{m\tau} + \frac{1}{m\tau} \right). \end{aligned} \quad (3.9)$$

Reorganizando a equação (3.9), temos a forma mais conhecida do modelo de

Nelson e Siegel.

$$R(m) = \beta_0 + \beta_1 \frac{[1 - \exp(-m\tau)]}{(m\tau)} + \beta_2 \left[\frac{[1 - \exp(-m\tau)]}{(m\tau)} - \exp(-m\tau) \right] + \epsilon. \quad (3.10)$$

Quando utilizamos o modelo definido na equação (3.10) para modelarmos a curva de taxa de juros, temos a presença de um erro de estimação (ϵ). Consequentemente, o erro estimado será:

$$\hat{\epsilon}_i = R_i - \hat{R}_i,$$

em que R_i é o rendimento real da maturidade i , \hat{R}_i é o rendimento ajustado obtido pelo modelo para a i -ésima maturidade.

O erro estimado é a diferença entre o valor observado para a taxa de juros e o valor ajustado pelo modelo. Dessa forma, quando o objetivo é modelar a ETTJ devemos incluir o erro de estimação no modelo.

Por suposição, dizemos que os ϵ_i são independentes, com média zero e variância constante.

É possível dizer que o modelo é uma constante adicionada à função ou polinômio de Laguerre. Polinômios de Laguerre são uma família de polinômios ortogonais em homenagem a Edmond Laguerre e aparecem na análise de soluções para equações diferenciais de segunda ordem. Detalhes sobre esses polinômios podem ser encontrados em Courant and Hilbert (1953, 1:93-97).

3.2.1 Interpretação dos Parâmetros

Vimos que o modelo tem quatro parâmetros e podemos dividi-los em duas classes: parâmetros de regressão (β_0, β_1 e β_2) e parâmetro de tempo (τ).

Diebold and Li (2006) identificam os parâmetros de regressão a fatores específicos responsáveis pela movimentação das curvas de taxas de juros. Assim,

por definição, o parâmetro β_0 é classificado como fator de nível, o parâmetro β_1 é classificado como fator de inclinação e o parâmetro β_2 como fator de curvatura. Na sequência apresentamos o motivo desses parâmetros serem associados aos fatores responsáveis pela movimentação das curvas de taxas de juros.

Parâmetros de Regressão

Os parâmetros de regressão podem ser associados aos fatores responsáveis pela movimentação das curvas de taxas de juros. Além disso, podemos mostrar que esses parâmetros estão associados ao prazo de vencimento dos títulos (curto, médio e longo prazos). Podemos dizer que o parâmetro β_0 é considerado um fator de longo prazo, o parâmetro β_1 é considerado um fator de curto prazo e por fim o parâmetro β_2 é considerado um fator de médio prazo, como podemos ver na sequência deste trabalho.

Para mostrar porque os parâmetros podem ser considerados fatores de curto, médio e longo prazos devemos observar as cargas fatoriais ou coeficientes dos parâmetros de regressão na equação do modelo. Podemos observar pela equação (3.10) que o parâmetro de tempo τ está presente nas cargas fatoriais dos parâmetros de regressão. Contudo, τ não é considerado para efeito de explicação por se tratar de uma constante que não influenciará na definição de um fator de curto, médio ou longo prazo.

A Figura 3.1 ilustra o comportamento das cargas fatoriais para as diferentes maturidades e assim podemos visualizar o motivo do parâmetro β_0 ser considerado fator de longo prazo, o parâmetro β_1 ser considerado fator de curto prazo, o parâmetro β_2 ser considerado fator de médio prazo e além disso, podemos notar o motivo da existência das restrições do modelo.

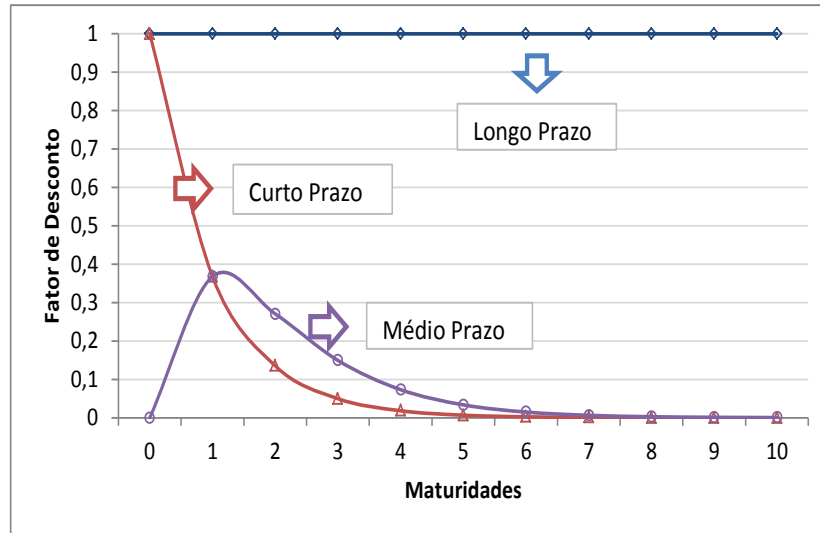


Figura 3.1: Cargas Fatoriais

1. Parâmetro β_0 : o termo que acompanha o parâmetro β_0 é uma constante, ou seja, um valor que não se altera independentemente dos prazos de vencimento do título. Como podemos ver pela Figura 3.1, esse coeficiente se mantém para as maturidades mais longas, enquanto que os demais para altas maturidades, apresentam valor tendendo a zero. Podemos mostrar esse fato matematicamente, para isso, devemos fazer m tendendo ao infinito ($m \rightarrow \infty$) na equação (3.10).

$$R(m) \lim_{m \rightarrow \infty} \beta_0 + \beta_1 \frac{[1 - \exp(-m\tau)]}{(m\tau)} + \beta_2 \left[\frac{[1 - \exp(-m\tau)]}{(m\tau)} - \exp(-m\tau) \right] = \beta_0.$$

Portanto, quando necessitamos de uma rápida estimativa para o parâmetro

β_0 , podemos utilizar o valor do rendimento de um título com vencimento mais longo.

Além desse parâmetro ser considerado um fator de longo prazo, podemos relacioná-lo ao nível da curva de taxa de juros, um dos fatores responsáveis pela movimentação da curva (Seção 2.2.3). Dessa forma, qualquer mudança de valor nesse parâmetro altera de forma igual os rendimentos, alterando assim o nível da curva de taxas de juros.

2. Parâmetro β_1 : o termo que acompanha esse parâmetro é $\exp(-m)$ que, para os diversos vencimentos, é monotonamente decrescente, começando em um e decaindo rapidamente para zero como podemos verificar pela Figura (3.1). Por assumir valores significativos para títulos de curto prazo de vencimento e aproximadamente zero para as demais maturidades, esse parâmetro é considerado um fator de curto prazo. Dessa forma, se β_1 se altera, os rendimentos de curto prazo se alteram e isso muda a inclinação da curva. Isso mostra que β_1 , está intimamente relacionado com a inclinação da curva definida como $R(0) - R(\infty) = \beta_1$.
3. Parâmetro β_2 : o termo que o acompanha é $m \exp(-m)$, ou seja, começa em zero, assume valores positivos para prazos médios e retorna a zero para vencimentos longos. Por assumir valores significativos apenas para títulos de médio prazo esse parâmetro é considerado um fator de médio prazo.

Esse parâmetro ainda está intimamente relacionado com a curvatura da curva de taxa de juros, que pode ser calculada da forma $2R(m^*) - R(\infty) - R(0)$, onde m^* representa uma maturidade média. Por exemplo, se estamos interessados no rendimento de um título com vencimento de 4 anos, devemos adotar como m^* o valor médio, neste caso $m^* = 2$. Como sabemos que $R(0) = \beta_1 + \beta_0$ e $R(\infty) = \beta_0$ então temos que $R(m^*) = \frac{\beta_2 + 2\beta_0 + \beta_1}{2}$. Logo, com

o aumento desse parâmetro, os rendimentos de curto e longo prazos não são influenciados, porém, os rendimentos de médio prazo serão, e isso implicará na alteração da curvatura da curva.

Parâmetro de Tempo

Além dos parâmetros de regressão, o modelo de Nelson-Siegel apresenta o parâmetro de tempo τ relacionado com a velocidade de decaimento da curva de taxa de juros, ou seja, a velocidade com que as taxas de juros tendem para seu valor de longo prazo. Valores pequenos desse parâmetro produzem um decaimento rápido, indicando um melhor ajuste para títulos de curto prazo e valores mais altos produzem um decaimento mais lento, indicando um melhor ajuste para os títulos com vencimentos mais longo.

Em geral, quando desejamos fazer a estimação desse parâmetro, quatro formas são as mais utilizadas no mercado financeiro.

1. Estabelecer um valor arbitrário qualquer.
2. Média dos parâmetros variáveis no tempo, que foram estimados ao longo de todo o período da amostra.
3. Estimar juntamente com os demais parâmetros.
4. Minimizar o erro médio quadrático (MSE) da previsão.

Em geral, como no mercado financeiro é necessário obter as informações de forma rápida o tratamento desse parâmetro é realizado de acordo com o primeiro critério. Diebold e Li(2006) mostram em um estudo, para dados americanos, que não há perda de generalidade quando assumimos que esse parâmetro é considerado constante, e isso pode ser explicado pelo fato dele apenas determinar o vencimento

para o qual o peso associado à curvatura é maximizado, porém em outros países esse fato não é assumido e a estimação desse parâmetro deve ser realizada.

Considerando que o parâmetro de tempo é constante, temos que a equação (3.10) que define o modelo de Nelson-Siegel se torna linear nos coeficientes e deste modo temos várias formas de estimação conhecidas e de fácil aplicação.

Uma visualização de como o parâmetro de tempo se ajusta melhor a títulos de curto ou longo prazos, dependendo se sua estimativa assumir valores pequenos ou grandes, pode ser encontrada no trabalho original publicado por Nelson e Siegel (Nelson and Siegel, 1987, 477:479).

Pela Figura (3.1) e pela definição do parâmetro de tempo, podemos entender o motivo das restrições do modelo. Inicialmente devemos saber que pelo fato de estarmos trabalhando com taxa de juros as estimativas para elas devem assumir valores positivos. Dessa forma, podemos explicar a restrição $\hat{\beta}_0 \geq 0$ pelo fato desse fator ser o único atuante nos títulos de longo prazo de vencimento. A restrição $\hat{\beta}_0 + \hat{\beta}_1 \geq 0$ é dada pois, como observamos, o parâmetro β_1 é considerado um fator de curto prazo e que atua conjuntamente com o parâmetro β_0 e como as taxas de juros devem ser positivas, essa restrição deve ser imposta. Por último temos a restrição $\hat{\tau} > 0$, em que suas estimativas devem ser positivas, isso ocorre porque os vencimentos medem o tempo.

3.3 Esboço da ETTJ pelas estimativas dos parâmetros

O modelo em estudo apresenta uma importante vantagem em sua aplicação: conhecendo as estimativas ou uma ideia grosseira acerca das estimativas dos parâmetros, podemos esboçar o formato que a curva de taxa de juros pode assumir (Normal, Invertida, Ondulada ou Plana) sem que seja necessário ajustar o modelo.

Supondo que as estimativas dos parâmetros β_1 e β_2 podem assumir quaisquer valores, teremos os seguintes formatos de curvas de taxas de juros.

- Estimativa do parâmetro β_1

$\hat{\beta}_1 < 0 \Rightarrow$ a curva para curto prazo de vencimento é crescente.

$\hat{\beta}_1 > 0 \Rightarrow$ a curva para curto prazo de vencimento é decrescente.

- Estimativa do parâmetro β_2

$\hat{\beta}_2 < 0 \Rightarrow$ a curva terá um formato de “U”.

$\hat{\beta}_2 > 0 \Rightarrow$ a curva terá um formato de “U invertido”.

$\hat{\beta}_2 \cong 0 \Rightarrow$ teremos uma função monótona crescente ou decrescente dependendo do valor da estimativa do parâmetro β_1 .

A Figura (3.2) apresenta um esboço do formato das curvas de taxas de juros, levando em conta apenas se as estimativas assumem valores positivos ou negativos para os parâmetros β_1 e β_2 .

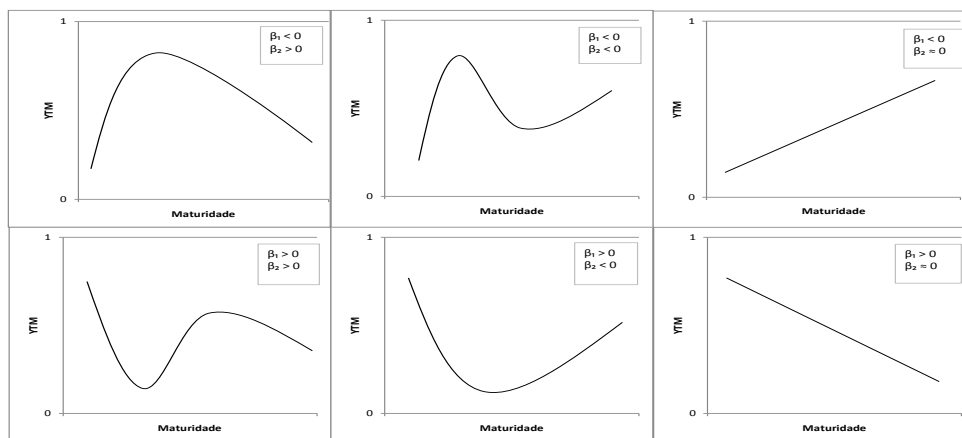


Figura 3.2: Ilustração do formato das curvas

Capítulo 4

Métodos de Estimação

Com o modelo de Nelson e Siegel definido no capítulo três, o objetivo agora é obtermos as estimativas pontuais para seus parâmetros para assim podermos construir uma curva de taxas de juros que melhor represente os dados de um estudo (versão estática).

Outros trabalhos referente ao modelo de Nelson e Siegel podem ser encontradas em Laurini e Hotta (2007), Guedes (2008), Laurini e Neto (2010), entre outros.

Para iniciarmos o processo de estimação é necessário definir qual será a técnica usada para estimar o parâmetro de tempo τ , para assim adotarmos técnicas para estimação dos parâmetros de regressão. Como Diebold e Li(2006) mostraram que não há perda de generalidade quando assumimos que τ é considerado constante, adotamos inicialmente, um valor fixo e arbitrário para esse parâmetro. De forma que o modelo passa a ser linear nos coeficientes β_0 , β_1 e β_2 .

No *software R* o pacote *YieldCurve*, oferece uma rotina para estimar os parâmetros do modelo de Nelson e Siegel. Contudo como veremos pelo exemplo a seguir, nem sempre as estimativas obtidas são boas, pois as restrições do modelo não são respeitadas.

A partir do pacote *YieldCurve* do *software R* foi gerado um conjunto de dados,

Tabela 4.1, à partir de valores assumidos para os parâmetros do modelo.

Tabela 4.1: Conjunto de Dados - Exemplo Software R

Maturidade	Taxa de Juros (%a.a.)	Maturidade	Taxa de Juros (%a.a.)
0,5	2,474428	30	3,083001
1	5,869792	31	2,951391
2	10,262153	32	2,827976
3	12,474859	33	2,712022
4	13,347093	34	2,602875
5	13,417521	35	2,499956
6	13,028611	36	2,402748
7	12,395321	37	2,310791
8	11,650219	38	2,223671
9	10,873019	39	2,141017
10	10,109851	40	2,062494
11	9,385772	41	1,987801
12	8,712834	42	1,916664
13	8,095266	43	1,848835
14	7,532751	44	1,78409
15	7,022486	45	1,722221
16	6,560451	46	1,663043
17	6,142168	47	1,606383
18	5,763145	48	1,552083
19	5,419119	49	1,5
20	5,106172	50	1,45
21	4,820778	51	1,401961
22	4,559806	52	1,355769
23	4,320497	53	1,311321
24	4,100435	54	1,268518
25	3,897508	55	1,227273
26	3,709873	56	1,1875
27	3,535925	57	1,149123
28	3,374256	58	1,112069
29	3,223641	59	1,076271
		60	1,041667

Pelo programa existente no *software R* as estimativas para os parâmetros do

modelo são:

Tabela 4.2: Estimativas dos Parâmetros - Exemplo Software R

β_0	β_1	β_2
-6,4087	17,6175	7,9182

Com base no resultado acima podemos notar que a restrição do modelo $\beta_0 \geq 0$, não é satisfeita.

Para incorporar as restrições na estimação dos parâmetros do modelo de Nelson e Siegel, apresentamos nesse capítulo os métodos de Mínimos Quadrados Ordinários (MQO) (Seção 4.1), Máxima Verossimilhança (MV) (Seção 4.2), detalhes sobre os métodos e as propriedades podem ser vistos em Graybill (2000) e em Casella & Berger (2001), e a abordagem bayesiana (Seção 4.3).

Sabemos que as restrições do modelo definidas na Seção 3.2 são restrições de desigualdade. Para utilizarmos os métodos de MQO e MV sob essas restrições, utilizamos o método proposto por Karush, Kuhn e Tucker, denominado condições KKT. Maiores detalhes sobre o método podem ser vistos em W. Karush (1939) e H. W. Kuhn & A. W. Tucker (1951). Outra referência sobre estimação com restrição pode ser encontrada em C. Gourieroux & A. Monfort (1995).

De modo geral, o método KKT consiste em maximizar ou minimizar uma função de interesse com certas restrições de desigualdade e/ou igualdade. Ou seja,

minimizar/maximizar $f(x)$, sujeito a:

$$g_i(x) \geq 0$$

$$h_j(x) = 0$$

em que a função $f(x)$ é a função que queremos minimizar ou maximizar, x é a variável de otimização, g_i são as restrições de desigualdade com ($i = 1, \dots, p$) e h_j são as restrições de igualdade com ($j = 1, \dots, l$).

Sendo que:

$$f: \mathfrak{R}^n \rightarrow \mathfrak{R}, g_i: \mathfrak{R}^n \rightarrow \mathfrak{R} \text{ e } h_j: \mathfrak{R}^n \rightarrow \mathfrak{R}.$$

Se essas funções forem continuamente diferenciáveis em algum ponto, então existem constantes μ_i , com $(i = 1, \dots, p)$ e λ_j , com $(j = 1, \dots, l)$, de tal forma que:

$$\nabla f(x) + \sum_{i=1}^p \mu_i \nabla g_i(x) + \sum_{j=1}^l \lambda_j \nabla h_j(x) = 0$$

em que ∇ é o gradiente (derivadas parciais da função), $f(x)$ é a função que queremos minimizar ou maximizar, μ_i e λ_j são variáveis auxiliares e $g_i(x)$ e $h_j(x)$ são restrições de desigualdade e igualdade, respectivamente, do modelo.

$$g_i(x) \geq 0, \text{ para todo } i = 1, \dots, p.$$

$$h_j(x) = 0, \text{ para todo } j = 1, \dots, l.$$

$$\mu_i \leq 0, \text{ para todo } i = 1, \dots, p.$$

$$\mu_i g_i(x) = 0, \text{ para todo } i = 1, \dots, p.$$

Através do cálculo dessas equações conseguimos encontrar as estimativas para os parâmetros.

Por simplicidade de notação denotaremos no modelo de Nelson e Siegel o coeficiente que acompanha o parâmetro β_1 por X_{1i} e o coeficiente que acompanha o parâmetro β_2 por X_{2i} , para efeitos de cálculos.

4.1 Mínimos Quadrados Ordinários

Quando realizamos a estimação dos parâmetros por Mínimos Quadrados Ordinários temos por objetivo minimizar a SQR, ou seja, $SQR(\beta) = \sum_{i=1}^n \hat{\epsilon}_i^2$ deve ser a menor possível, em que $\hat{\epsilon}$ é o erro estimado conforme definido na Seção (3.2). A vantagem apresentada pela estimação via MQO é a não necessidade de suposição de hipóteses sobre a distribuição das variáveis do modelo.

Como vimos, nosso objetivo é minizar a SQR, sujeito as restrições $\beta_0 \geq 0$ e $\beta_0 + \beta_1 \geq 0$, ou seja:

$$\min \sum_{i=1}^n [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}]^2. \quad (4.1)$$

sujeito a:

$$\begin{cases} \beta_0 \geq 0 \\ \beta_0 + \beta_1 \geq 0 \end{cases}$$

Como podemos observar, não temos restrições de igualdade no modelo, ou seja, a parte da equação referente as restrições de igualdade é 0.

Para obtermos as estimativas para os parâmetros pelo método dos MQO devemos estabelecer as condições KKT de primeira ordem, que são dadas por:

(i)

$$-2 \sum_{i=1}^n [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] + \mu_1 + \mu_2 = 0$$

(ii)

$$-2 \sum_{i=1}^n X_{1i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] + \mu_2 = 0$$

(iii)

$$-2 \sum_{i=1}^n X_{2i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] = 0$$

(iv)

$$\beta_0 \mu_1 = 0$$

(v)

$$(\beta_0 + \beta_1) \mu_2 = 0$$

(vi)

$$\mu_1 \leq 0, \mu_2 \leq 0$$

(vii)

$$g_1 \geq 0, g_2 \geq 0$$

Com as condições KKT definidas acima, podemos ter quatro soluções:

Caso1: $\mu_1 = 0$ e $\mu_2 = 0$.

As restrições $\beta_0 \geq 0$ e $\beta_0 + \beta_1 \geq 0$ não são satisfeitas, pois $\hat{\beta}_0 \leq 0$ e $\hat{\beta}_0 + \hat{\beta}_1 \leq 0$.

Caso2: $\mu_1 = 0$ e $\mu_2 < 0$.

Temos que por (i) e (ii):

$$\hat{\mu}_2 = 2 \sum_{i=1}^n [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}]$$

$$\hat{\mu}_2 = 2 \sum_{i=1}^n X_{1i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}]$$

Como temos diferentes expressões para determinarmos $\hat{\mu}_2$ que resultam em valores diferentes, não podemos considerar esse caso.

Caso3: $\mu_1 < 0$ e $\mu_2 = 0$.

Então por (iv) temos que $\hat{\beta}_0 = 0$ e por (v) $\hat{\beta}_0 = -\hat{\beta}_1 = 0$.

Verificamos em (ii) e (iii) que:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{1i}R_i}{\sum_{i=1}^n X_{1i}X_{2i}}$$

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{2i}R_i}{\sum_{i=1}^n X_{2i}X_{2i}}$$

Como temos valores diferentes para $\hat{\beta}_2$ não podemos considerar esse caso.

Caso4: $\mu_1 < 0$ e $\mu_2 < 0$.

Então por (iv) temos que $\hat{\beta}_0 = 0$ e por (v) $\hat{\beta}_0 = -\hat{\beta}_1 = 0$.

Por (iii) temos que:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{2i}R_i}{\sum_{i=1}^n X_{2i}X_{2i}}$$

Para o conjunto de dados definido na Tabela 2.1, temos que $\hat{\beta}_2 = 42,48616$.

Por (ii) temos que:

$$\begin{aligned}\hat{\mu}_2 &= 2 \sum_{i=1}^n X_{1i}R_i + 2\hat{\beta}_2 \sum_{i=1}^n X_{1i}X_{2i} \\ \hat{\mu}_2 &= -1,663657\end{aligned}$$

E por fim, por (i) temos:

$$\begin{aligned}\hat{\mu}_1 &= 2 \sum_{i=1}^n R_i + 2\hat{\beta}_2 \sum_{i=1}^n X_{2i} - mu_2 \\ \hat{\mu}_1 &= -35,1076\end{aligned}$$

Ou seja, temos que todas as condições KKT são satisfeitas.

Dessa forma, para esse caso, as estimativas para os parâmetros do modelo são dadas por:

Tabela 4.3: Estimativas dos Parâmetros - Exemplo MQO com restrição

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0	0	42,48616

Podemos verificar pela Figura 4.1 o comportamento da curva ajustada com relação a curva real.

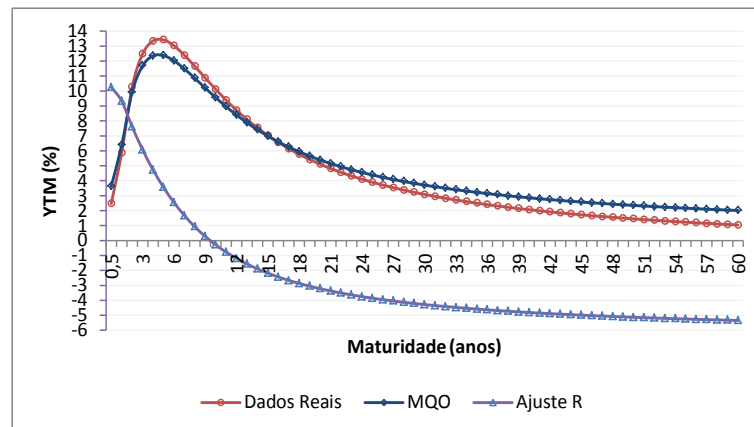


Figura 4.1: Curva de Taxas de Juros - Exemplo

Pela Figura 4.1 podemos notar que o ajuste realizado pelo *software R* apresenta uma curva em que a maioria dos valores são negativos, o que nos indica um erro no ajuste. Observamos também que a distância dessa curva para a curva real é

muito grande. Por outro lado, a curva construída pelo método MQO com restrição apresenta seu comportamento similar ao da curva real, indicando um bom ajuste por esse método.

4.2 Máxima Verossimilhança

Diferentemente do método dos Mínimos Quadrados Ordinários, em que não era necessário supor distribuição para as variáveis, para usar o método da Máxima Verossimilhança essa suposição deverá ser utilizada. O objetivo desse método é encontrar o valor que maximiza a função de verossimilhança.

Assumindo o vetor de erros (ϵ) com distribuição $N(0, \sigma^2 I)$, onde I é a matriz identidade, então R também tem distribuição normal, com média dada por:

$$E(R) = E(X\beta + \epsilon) = E(X\beta) + E(\epsilon) = X\beta,$$

e variância:

$$Var(R) = Var(X\beta + \epsilon) = Var(X\beta) + Var(\epsilon) = \sigma^2 I.$$

Portanto, temos que R apresenta distribuição $N(X\beta, \sigma^2 I)$. Sabemos que quando assumimos que a taxa apresenta uma distribuição Gaussiana o resultado apresentado é o mesmo encontrado através do método dos MQO.

Definida a distribuição, a função de verossimilhança é dada por.

$$L(\beta, \sigma^2 | R, X) = \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right]^n \exp \left[\frac{-1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right]$$

$$L(\beta, \sigma^2 | R, X) = [2\pi\sigma^2]^{-n/2} \exp \left[\frac{-1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right].$$

Aplicando o logaritmo na função de verossimilhança $\ln(L(\beta, \sigma^2|R, X))$ temos:

$$l(\beta, \sigma^2, R, X) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

Como vimos nosso objetivo é maximizar a função de verossimilhança ou a função log-verossimilhança, sujeito as restrições $\beta_0 \geq 0$ e $\beta_0 + \beta_1 \geq 0$.

Estabelecendo as condições KKT de primeira ordem.

(i)

$$\frac{2}{2\sigma^2} \sum_{i=1}^n [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] + \mu_1 + \mu_2 = 0$$

(ii)

$$\frac{2}{2\sigma^2} \sum_{i=1}^n X_{1i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] + \mu_2 = 0$$

(iii)

$$\frac{2}{2\sigma^2} \sum_{i=1}^n X_{2i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}] = 0$$

(iv)

$$\frac{-n}{2\sigma^2} + \frac{2}{(2\sigma^2)^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 x_1 - \beta_2 x_2)^2 = 0$$

(v)

$$\beta_0 \mu_1 = 0$$

(vi)

$$(\beta_0 + \beta_1) \mu_2 = 0$$

(vii)

$$\mu_1 \leq 0, \mu_2 \leq 0$$

(viii)

$$g_1 \geq 0, g_2 \geq 0$$

Com as condições KKT definidas acima podemos ter quatro soluções:

Caso1: $\mu_1 = 0$ e $\mu_2 = 0$.

As restrições sobre $\beta_0 \geq 0$ e $\beta_0 + \beta_1 \geq 0$ não são satisfeitas, pois $\hat{\beta}_0 \leq 0$ e $\hat{\beta}_0 + \hat{\beta}_1 \leq 0$.

Caso2: $\mu_1 = 0$ e $\mu_2 < 0$.

Temos que por (i) e (ii):

$$\hat{\mu}_2 = 2 \sum_{i=1}^n [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}]$$

$$\hat{\mu}_2 = 2 \sum_{i=1}^n X_{1i} [R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}]$$

Como temos diferentes expressões para determinarmos $\hat{\mu}_2$ que resultam em valores diferentes, não podemos considerar esse caso.

Caso3: $\mu_1 < 0$ e $\mu_2 = 0$.

Então por (v) temos que $\hat{\beta}_0 = 0$ e por (vi) $\hat{\beta}_0 = -\hat{\beta}_1 = 0$.

Verificamos em (ii) e (iii) que:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{1i} R_i}{\sum_{i=1}^n X_{1i} X_{2i}}$$

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{2i} R_i}{\sum_{i=1}^n X_{2i} X_{2i}}$$

Como temos valores diferentes para $\hat{\beta}_2$ não podemos considerar esse caso.

Caso4: $\mu_1 < 0$ e $\mu_2 < 0$.

Então por (v) temos que $\hat{\beta}_0 = 0$ e por (vi) $\hat{\beta}_0 = -\hat{\beta}_1 = 0$.

Por (iii) temos que:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_{2i}R_i}{\sum_{i=1}^n X_{2i}X_{2i}}$$

Para o conjunto de dados definido na Tabela 2.1, temos que $\hat{\beta}_2 = 42,48616$.

Por (iv) temos que:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n [R_i - \beta_2 X_{2i}]^2}{n}$$

Com isso, temos que $\hat{\sigma}^2 = 1452,061$.

Por (ii) temos que:

$$\begin{aligned}\hat{\mu}_2 &= 2 \sum_{i=1}^n X_{1i}R_i + 2\beta_2 \sum_{i=1}^n X_{1i}X_{2i} \\ \hat{\mu}_2 &= -1,663657\end{aligned}$$

E por fim, por (i) temos:

$$\begin{aligned}\hat{\mu}_1 &= 2 \sum_{i=1}^n R_i + 2\beta_2 \sum_{i=1}^n X_{2i} - mu_2 \\ \hat{\mu}_1 &= -35,1076\end{aligned}$$

Desta forma, todas as condições KKT são satisfeitas.

Verificamos assim que as estimativas para os parâmetros do modelo via MV

com restrição são as mesmas encontradas pelo método MQO com restrição. Desta forma, a Figura 4.1 representa tanto a estimativa da curva de taxas de juros obtida pelo método de MQO com restrição como a obtida pelo método da MV com restrição.

Uma observação sobre as condições KKT é que, independente do conjunto de dados que queremos analisar, não podemos seguir a análise no caso em que ($\mu_1 = 0$ e $\mu_2 < 0$) e no caso em que ($\mu_1 < 0$ e $\mu_2 = 0$). Por outro lado, dependendo do conjunto de dados, das outras duas possibilidades de resolução, apenas uma será aceita pelas condições KKT. Portanto, sempre que estamos utilizando as condições KKT devemos verificar o valor das estimativas para os parâmetros para os dois casos e assim obter as estimativas corretas para os parâmetros.

4.3 Inferência Bayesiana - τ fixo

A origem da Inferência Bayesiana aconteceu com a publicação do artigo *An Essay Towards Solving a Problem in the Doctrine of Chances*, a título póstumo, de autoria do Reverendo Thomas Bayes, publicado por Richard Price, no qual foi apresentado o teorema denominado Teorema de Bayes (Bayes, 1763), fundamentado em probabilidades condicionais (Mood et al., 1974). O teorema de Bayes é uma regra para atualização da distribuição a posteriori e que independente da ordem em que processamos as informações, o resultando se mantém sempre o mesmo.

A inferência bayesiana ganhou espaço com o surgimento de novos recursos computacionais e algoritmos computacionais como os algoritmos *Monte Carlo Markov Chain* (MCMC).

O objetivo da inferência bayesiana está na obtenção da distribuição a posteriori. Essa distribuição representa todo o conhecimento existente sobre o problema

investigado. E toda pergunta a respeito do modelo será respondida a partir de sua análise. Qualquer aspecto do parâmetro de interesse, como a média, mediana, moda, variância, percentis, fazer previsões para outras observações, etc, são obtidos a partir da análise dessa distribuição.

A distribuição a posteriori é proporcional ao produto entre a função de verossimilhança e a distribuição a priori.

$$\pi(\text{Parâmetros}|\text{Dados}) \propto L(\text{Parâmetros}, \text{Dados})\pi(\text{Parâmetros}). \quad (4.2)$$

em que $L(\text{Parâmetros}, \text{Dados})$ é a função de verossimilhança e $\pi(\text{Parâmetros})$ é a distribuição a priori.

Maiores detalhes sobre a inferência bayesiana, como por exemplo, como definir a distribuição a priori, verossimilhança e os métodos computacionais para gerar a amostra da posteriori (Métodos Computacionais MCMC) podem ser vistos em Ehlers (2005), Robert & Casella (1999), Paulino & Turkman & Murteira (2003) e Gamerman (1997).

No capítulo três, vimos que o modelo de Nelson e Siegel apresenta em sua definição algumas restrições que devem ser impostas quando vamos determinar as estimativas dos parâmetros. A inferência bayesiana nos permite incorporar essas restrições através da definição das prioris adequadas.

Assumindo que os rendimentos da taxa de juros apresentam distribuição normal e adotando como priori para cada parâmetro as seguintes distribuições, temos:

1. Priori para β_0 :

Como esse parâmetro deve ser não-negativo, adotamos como priori uma distribuição Gama(a,b), onde devemos escolher valores para os hiperparâmetros a e b .

2. Priori para β_1 :

Como temos a restrição que $\beta_0 + \beta_1 \geq 0$, então temos que $\beta_1 \geq -\beta_0$. Pela Figura (4.1), podemos visualizar a área do sub-espço paramétrico em que a restrição é válida.

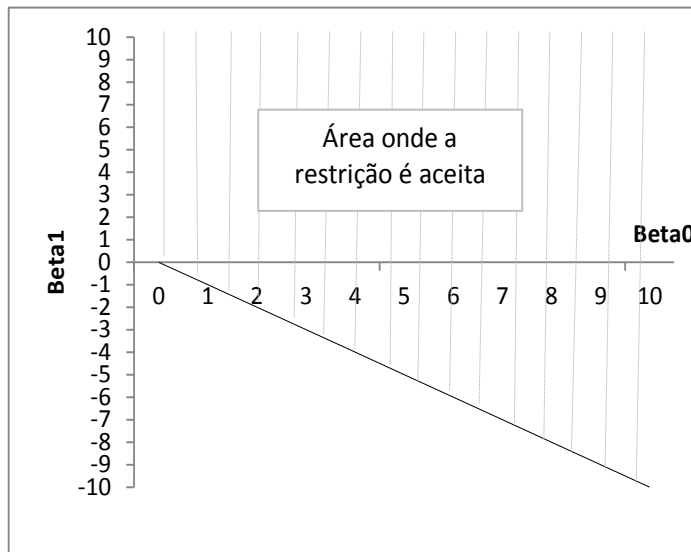


Figura 4.2: Restrição

Dessa forma, podemos adotar como priori para $\beta_1|\beta_0$ uma Normal Truncada (μ_1, σ_1^2) , em que o intervalo de variação para o parâmetro é $(-\beta_0, \infty)$.

3. Priori para β_2 :

Como não temos restrições para a estimativa deste parâmetro, utilizaremos como distribuição a priori a distribuição Normal (μ_2, σ_2^2)

4. Priori para σ^2 :

Admitimos a priori proporcional a $\frac{1}{\sigma^2}$.

Essas prioris foram escolhidas apenas para ilustrar como as restrições são incor-

poradas no processo inferencial. Outras distribuições que respeitem as restrições sobre os parâmetros de regressão β_0 e β_1 discutidas acima podem ser utilizadas.

As prioris são dadas, respectivamente por:

$$\pi(\beta_0) = \frac{b^a}{\Gamma(a)} \beta_0^{(a-1)} \exp(-b\beta_0). \quad (4.3)$$

$$\pi(\beta_1|\beta_0) = \frac{1}{1 - \Phi(\beta_0)} \left(\frac{1}{2\pi\sigma_1^2} \right)^{-1/2} \exp\left(-\frac{1}{2\sigma_1^2}(\beta_1 - \mu_1)^2\right). \quad (4.4)$$

em que $\Phi(\beta_0)$ se refere a função de distribuição normal padrão avaliada em β_0

$$\pi(\beta_2) = \left[\frac{1}{2\pi\sigma^2} \right]^{-1/2} \exp\left[-\frac{1}{2\sigma^2}(\beta_2 - \mu_2)^2\right]. \quad (4.5)$$

$$\pi(\sigma^2) \propto \frac{1}{\sigma^2}. \quad (4.6)$$

A posteriori é proporcional ao produto das prioris com a verossimilhança:

$$\begin{aligned} \pi(\beta_0, \beta_1, \beta_2, \sigma^2|D) \propto & \left[\frac{1}{1 - \Phi(\beta_0)} \right] \frac{b^a}{\Gamma(a)} \beta_0^{(a-1)} \left(\frac{1}{2\pi\sigma_1^2} \right)^{1/2} \frac{1}{\sigma^2} \left(\frac{1}{2\pi\sigma_2^2} \right)^{1/2} \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \\ & \exp\left[-\frac{1}{2\sigma_1^2}(\beta_1 - \mu_1)^2 - b\beta_0 - \frac{1}{2\sigma_2^2}(\beta_2 - \mu_2)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2\right]. \end{aligned} \quad (4.7)$$

Para calcular as quantidades de interesse dessa distribuição utilizaremos os procedimentos de Monte Carlo e para isso necessitamos de amostras i.i.d. dessa distribuição.

Dada a posteriori, encontramos assim as condicionais completas para cada parâmetro, que são proporcionais às expressões a seguir:

1. Condicional para σ^2

$$\pi(\sigma^2|\beta, D) \propto \left[\frac{1}{\sigma^2} \right]^{\left(\frac{n}{2}+1\right)} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right]^2. \quad (4.8)$$

Portanto, σ^2 , dados os demais parâmetros e os dados tem:

$$\text{Distribuição Gama Inversa} \left[\frac{n}{2}, \frac{\sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{2} \right].$$

2. Condicional para β_0

$$\pi(\beta_0|\beta_1, \beta_2, \sigma^2, D) \propto \left[\frac{1}{1 - \phi(\beta_0)} \right] \beta_0^{(a-1)} \exp \left[-b\beta_0 - \frac{1}{2\sigma^2} \sum (R_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2)^2 \right]. \quad (4.9)$$

O parâmetro β_0 , dados os demais parâmetros e os dados não tem distribuição conhecida.

3. Condicional para β_1

$$\pi(\beta_1|\beta_0, \beta_2, \sigma^2, D) \propto \exp \left[-\frac{1}{2\sigma_1^2} (\beta_1 - \mu_1)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right]^2 \quad (4.10)$$

O parâmetro β_1 dados os demais parâmetros e os dados tem:

$$\text{Distribuição Normal} \left[\left(\frac{\sigma^2 \mu_1 + \sigma_1^2 \sum_{i=1}^n X_1 (R_i - \beta_0 - \beta_2 X_2)}{\sigma^2 + \sigma_1^2 \sum_{i=1}^n X_1^2} \right); \left(\frac{\sigma^2 \sigma_1^2}{\sigma^2 + \sigma_1^2 \sum_{i=1}^n X_1^2} \right) \right].$$

4. Condicional para β_2

$$\pi(\beta_2|\beta_0, \beta_1, \sigma^2, D) \propto \exp \left[-\frac{1}{2\sigma_2^2} (\beta_2 - \mu_2)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right]^2 \quad (4.11)$$

O parâmetro β_2 dados os demais parâmetros e os dados tem:

Distribuição Normal $\left[\left(\frac{\sigma^2 \mu_2 + \sigma_2^2 \sum_{i=1}^n X_2 (R_i - \beta_0 - \beta_1 X_1)}{\sigma^2 + \sigma_2^2 \sum_{i=1}^n X_2^2} \right); \left(\frac{\sigma^2 \sigma_2^2}{\sigma^2 + \sigma_2^2 \sum_{i=1}^n X_2^2} \right) \right]$.

Pelas condicionais obtidas, verificamos que para o parâmetro β_0 não temos uma distribuição conhecida, enquanto que para todos os demais parâmetros as distribuições condicionais são conhecidas.

4.3.1 Estudo de Simulação

Para ilustrar o procedimento de inferência bayesiana descrito na Seção anterior, realizaremos um estudo de simulação, no qual temos conhecimento dos verdadeiros valores dos parâmetros. Para todos os parâmetros com distribuição conhecida utilizamos o amostrador de Gibbs para gerar a amostra da distribuição a posteriori e para o parâmetro β_0 com a distribuição não conhecida utilizamos o método Metropolis-Hastings.

O estudo de simulação proposto consiste em avaliar a estimativa pontual dos parâmetros do modelo (100 replicações) com diferentes tamanhos de amostras, para isso consideramos tamanhos de amostras (vencimentos) $m = 5, 10, 30$ e 60 . Já para os parâmetros utilizamos 5 casos diferentes:

Caso 1: $\beta_0 = 1, \beta_1 = 5, \beta_2 = 6$ e $\sigma^2 = 0.01$

Caso 2: $\beta_0 = 0.1, \beta_1 = 3.5, \beta_2 = 3$ e $\sigma^2 = 0.05$

Caso 3: $\beta_0 = 3, \beta_1 = 1, \beta_2 = 1.5$ e $\sigma^2 = 0.001$

Caso 4: $\beta_0 = 5, \beta_1 = -2, \beta_2 = 1.5$ e $\sigma^2 = 1$

Caso 5: $\beta_0 = 1, \beta_1 = 0.5, \beta_2 = -2$ e $\sigma^2 = 2$

Realizamos 1000 iterações, com *burn-in* de 1000 e saltos de 150, obtendo assim, uma amostra final de 151000 valores gerados da posteriori. Definimos ainda

os valores para os hiperparâmetros das distribuições a priori: para β_0 que apresenta distribuição Gama(a,b), adotamos a=2 e b=20; para $\beta_1 \sim$ Normal Truncada (μ_1, σ_1^2) , adotamos os valores $\mu_1=0$ e $\sigma_1^2=100$; para $\beta_2 \sim$ Normal (μ_2, σ_2^2) adotamos os valores $\mu_2=0$ e $\sigma_2^2=100$. Além disso, foram calculadas para cada caso 100 estimativas dos parâmetros do modelo e para efeito de comparação será utilizado a média das estimativas encontradas como estimativa pontual.

Para o primeiro caso, calculamos as estatísticas descritivas (mínimo, média, máximo, variância, intervalos HPD¹ de 95% de confiança, medidas de erros (Erro Absoluto (EA), Erro Quadrático Médio (EQM), Raíz Erro Quadrático Médio (REQM), Erro Relativo Médio (ERM)) de acordo com as 100 estimativas dos parâmetros encontradas.

As estatísticas descritivas para o parâmetro β_0 podem ser observadas na tabela a seguir.

Tabela 4.4: Estimativas β_0 - Caso 1 - Simulação com τ fixo

m	Mín	Média	Máx	Var	Int. HPD 95%	EA	EQM	REQM	ERM
5	0,7253	0,8379	0,9333	0,0014	[0,7253;0,8966]	2,8629	0,0014	0,0375	1,5814
10	0,4998	0,9514	1,2930	0,0303	[0,5781;1,2421]	13,5437	0,03	0,1732	96,9086
30	0,8839	0,9983	1,1280	0,0026	[0,8924;1,0949]	4,1063	0,0027	0,0516	1,4646
60	0,9564	1	1,0590	0,0005	[0,9564;1,0435]	1,8648	0,0005	0,0230	1,4216

Notamos que para todos os casos temos que as estimativas, para o parâmetro β_0 , em média, são bem próximas do valor real do parâmetro ($\beta_0 = 1$). Além disso, podemos verificar que conforme aumentamos o número de vencimentos, mais precisas tornamos as estimativas encontradas. Pelas medidas dos erros notamos valores pequenos, o que indica que as estimativas estão bem próximas do valor real usado na simulação.

¹HPD do inglês Highest Posterior Density, ou seja, região de maior densidade a posteriori.

Na sequência temos as estatísticas descritivas para o parâmetro β_1 , que tem valor real 5.

Tabela 4.5: Estimativas β_1 - Caso 1 - Simulação com τ fixo

m	Mín	Média	Máx	Var	Int. HPD 95%	EA	EQM	REQM	ERM
5	4,751	5,082	5,446	0,0209	[4,8054;5,3647]	11,5048	0,0207	0,1438	0,8477
10	4,692	5,009	5,480	0,0263	[4,7518;5,3260]	12,8340	0,0260	0,1613	0,8692
30	4,7	4,999	5,275	0,0176	[4,7283;5,2221]	10,7904	0,0174	0,1319	0,7875
60	4,715	5,003	5,332	0,0144	[4,7359;5,1989]	9,7743	0,0143	0,1196	1,0502

Novamente, verificamos que as estimativas encontradas estão próximas ao valor real e a medida que aumentamos o número de vencimentos mais precisas se tornam as estimativas. Além disso, temos que as medidas novamente apresentam valores pequenos.

Para o parâmetro β_2 , com valor real 6, as estatísticas descritivas para as 100 estimativas são:

Tabela 4.6: Estimativas β_2 - Caso 1 - Simulação com τ fixo

m	Mín	Média	Máx	Var	Int. HPD 95%	EA	EQM	REQM	ERM
5	5,738	6,459	7,040	0,0731	[5,9684;7,0397]	21,6615	0,0724	0,2691	1,0328
10	4,979	6,145	7,363	0,3211	[5,1274;7,3395]	45,2494	0,3179	0,5638	1,0341
30	5,1	5,998	6,715	0,1334	[5,3224;6,7148]	29,6517	0,1320	0,3633	1,1694
60	5,369	6,008	6,630	0,0705	[5,4051;6,4813]	20,7955	0,0698	0,2641	1,7929

Assim como nos casos anteriores verificamos que as estimativas estão bem próximas ao valor real do o parâmetro e os erros pequenos.

Por fim, para o parâmetro de dispersão σ^2 com valor real de 0,01, temos as seguintes estatísticas descritivas:

Tabela 4.7: Estimativas σ^2 - Caso 1 - Simulação com τ fixo

m	Mín	Média	Máx	Var	Int. HPD 95%	EA	EQM	REQM	ERM
5	0,0013	0,0130	0,03524	$\cong 0$	[0,0031;0,0272]	0,5003	$\cong 0$	0,0065	0,8958
10	0,0034	0,0128	0,0374	$\cong 0$	[0,0035;0,0285]	0,4951	$\cong 0$	0,0065	0,9859
30	0,0051	0,0102	0,0167	$\cong 0$	[0,0055;0,0155]	0,2175	$\cong 0$	0,0027	0,8911
60	0,0064	0,0101	0,0141	$\cong 0$	[0,0076;0,0137]	0,1377	$\cong 0$	0,0017	1,0556

Novamente, podemos observar que a variabilidade das estimativas são pequenas e que os valores obtidos são bem próximos do valor médio que é aproximadamente igual ao valor real utilizado para a simulação. Pelos erros, temos que as estimativas encontradas são praticamente todas iguais ao valor real usado para a simulação.

Para os demais casos (Caso 2, Caso 3, Caso 4 e Caso 5), conforme tabela abaixo, também encontramos as estimativas para cada um dos parâmetros.

Tabela 4.8: Estimativas demais casos - Simulação com τ fixo

Caso	m	β_0		β_1		β_2		σ^2	
		Média	Var	Média	Var	Média	Var	Média	Var
2	5	0,1074	0,0003	3,4845	0,1357	3,0493	0,7867	0,1409	0,0101
2	10	0,1034	0,0002	3,5003	0,1060	3,0543	0,7921	0,0715	0,0019
2	30	0,1000	0,0011	3,5030	0,0777	3,0372	0,2453	0,0533	0,0002
2	60	0,0952	0,0010	3,4931	0,0773	3,0327	0,2439	0,0521	0,0001
3	5	2,9323	0,8231	1,0458	0,5539	1,3388	0,4771	0,0097	0,0011
3	10	2,8417	0,0206	1,0823	0,0077	1,7395	0,2800	0,0023	$\cong 0$
3	30	3	0,0003	1	0,0017	1,5007	0,0122	0,0011	$\cong 0$
3	60	3	$\cong 0$	1	0,0022	1,5	$\cong 0$	0,0010	$\cong 0$
4	5	4,9523	0,0231	-1,9305	2,0090	1,3885	5,9454	1,1872	3,4270
4	10	4,9876	0,0013	-1,9606	1,7730	1,6132	5,7924	1,0493	0,5106
4	30	5	$\cong 0$	-1,9616	2,0088	1,4797	4,8025	1,0384	0,0845
4	60	5	$\cong 0$	-1,9876	1,4198	1,5515	3,4612	1,0212	0,0301
5	5	0,9876	0,2378	0,5037	5,6852	-1,7783	5,3666	3,2825	7,9565
5	10	0,9912	0,0134	0,5447	3,3608	-1,9122	4,0828	2,5641	1,4978
5	30	0,9993	0,0001	0,5355	2,0141	-1,9334	2,0819	2,1022	0,3709
5	60	1	$\cong 0$	0,5319	1,3390	-1,9535	1,5891	2,0191	0,1477

Pelos resultados apresentados na Tabela 4.8, podemos verificar que as médias

das estimativas de seus respectivos parâmetros são aproximadamente iguais aos seus valores reais utilizados na simulação. Além disso, verificamos uma baixa variabilidade nas estimativas dos parâmetros.

Para confirmarmos a conclusão acima, uma nova simulação, com outros valores para os hiperparâmetros foi realizada. Novamente os resultados obtidos mostraram que as estimativas dos parâmetros estavam bem próximas de seus valores reais e que conforme aumentamos o número de informações no conjunto de dados, aproximamos ainda mais as estimativas do valor real do parâmetro. Desta forma, comprovamos mais uma vez a eficácia do método

4.4 Inferência Bayesiana - τ estimado

Até o momento, vimos formas de estimativas dos parâmetros do modelo apenas considerando o parâmetro de tempo τ fixo. Contudo, como vimos no capítulo 3, existem outras formas para estimação desse parâmetro, sendo uma delas a estimação conjunta com os parâmetros de regressão. Se deixarmos de considerar o parâmetro τ fixo o modelo que antes era considerado linear nos parâmetros, passa a ser não linear. Assim, para obtermos as estimativas dos parâmetros utilizaremos a abordagem bayesiana que incorpora as restrições do modelo na determinação das distribuições a priori.

Como vimos no capítulo 3, o parâmetro de tempo τ é não negativo. Dessa forma, atribuímos como priori para ele a distribuição Gama Inversa (c, d), em que c e d são os hiperparâmetros da distribuição. Além disso, utilizaremos para os demais parâmetros as distribuições definidas na Seção 4.3.

A escolha da distribuição Gama Inversa para a priori de τ foi proposta após, inicialmente, testarmos o uso da distribuição Gama, porém os resultados da simulação não foram satisfatórios. Outras distribuições, tais como a Log-Normal e

a Uniforme também são utilizadas em outros estudos e normalmente empregadas quando trabalhamos com a versão dinâmica do modelo.

Assim, como no caso em que o parâmetro de tempo era considerado fixo, devemos encontrar a distribuição a posteriori. A partir da amostra obtida por esta distribuição, encontramos as estimativas dos parâmetros do modelo. Sabemos que a distribuição a posteriori é proporcional ao produto da função de verossimilhança e a distribuição a priori (equação (4.2)). Assim, temos que à posteriori é dada por:

$$\pi(\beta_0, \beta_1, \beta_2, \tau, \sigma^2 | D) \propto \left[\frac{1}{1 - \phi(\beta_0)} \right] \frac{b^a}{\Gamma(a)} \beta_0^{(a-1)} \frac{d^c}{\Gamma(c)} \tau^{(-c+1)} \left(\frac{1}{2\pi\sigma_1^2} \right)^{1/2} \frac{1}{\sigma^2} \left(\frac{1}{2\pi\sigma_2^2} \right)^{1/2} \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left[-\frac{1}{2\sigma_1^2} (\beta_1 - \mu_1)^2 - b\beta_0 - \frac{d}{\tau} - \frac{1}{2\sigma_2^2} (\beta_2 - \mu_2)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (R_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \right] \quad (4.12)$$

Definida a distribuição a posteriori, nosso objetivo agora está em encontrarmos as condicionais completas para cada parâmetro e assim obter amostras independentes e identicamente distribuídas do vetor de parâmetros usando procedimentos de MCMC.

Para os parâmetros de regressão ($\beta_0, \beta_1, \beta_2$) e para o parâmetro de dispersão (σ^2) as condicionais são as mesmas obtidas pelas equações (4.8), (4.9), (4.10) e (4.11). Para o parâmetro de tempo (τ) a condicional é dada por:

$$\pi(\tau | \beta_0, \beta_1, \beta_2, \sigma^2, D) \propto \tau^{(-c+1)} \exp \left[-\frac{d}{\tau} - \frac{1}{2\sigma^2} \sum (R_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2)^2 \right]. \quad (4.13)$$

Portanto o parâmetro τ dado os demais parâmetros e os dados, não tem uma distribuição conhecida.

4.4.1 Estudo de Simulação

Para o estudo de simulação, adotaremos como estimativa pontual para os parâmetros a média a posteriori estimada através da amostra gerada da posteriori conjunta.

No estudo de simulação temos o conhecimento dos verdadeiros valores dos parâmetros. Para obtermos a amostra a posteriori, para os parâmetros β_0 e τ , dado que os mesmos não apresentam distribuições conhecidas, utilizaremos o método MCMC denominado Metropolis-Hastings e para os demais utilizaremos o amostrador de Gibbs

Realizamos 1000 iterações, com *burn-in* de 1000 e saltos de 10, obtendo assim, uma amostra final de 11000 valores gerados da posteriori. Para cada uma das simulações foram calculadas 100 replicações, ou seja, obtemos 100 estimativas para o parâmetro e através destas encontramos a estimativa pontual para o parâmetro. Definimos ainda os valores para os hiperparâmetros das distribuições a priori e adotamos 60 vencimentos.

- $\beta_0 \sim \text{Gama}(1,10)$.
- $\beta_1 \sim \text{Normal Truncada}(-10,100)$.
- $\beta_2 \sim \text{Normal}(0,100)$.
- σ^2 a priori proporcional a $1/\sigma^2$.
- $\tau \sim \text{Gama Inversa}(1,2)$.

Na Tabela (4.9) estão os resultados obtidos na simulação.

Pelos resultados podemos concluir que o método apresenta resultados satisfatórios, ou seja, estima com precisão os parâmetros do modelo. Dessa forma,

Tabela 4.9: Estudo Simulação com τ estimado

β_0		β_1		β_2		σ^2		τ	
Real	Média	Real	Média	Real	Média	Real	Média	Real	Média
0,5	0,4886	1	1,0264	7	7,0123	0,1	0,0991	0,7	0,7425
4	3,9997	2	1,9894	6	6,0461	0,05	0,0498	0,5	0,5081
3	2,9980	3	3,0231	2,5	2,5785	0,03	0,0307	3	2,8695
1,5	1,5050	4	4,1009	-2	-2,0666	0,01	0,0098	2	2,0159
2	2,0525	-1	-0,9923	0,5	0,4881	0,5	0,5118	0,3	0,3585

podemos utilizar esse método de estimação com uma maior precisão do que atribuindo um valor arbitrário para o parâmetro de tempo (τ).

Capítulo 5

Aplicação em Dados Reais

Neste capítulo estimamos os parâmetros do modelo de Nelson e Siegel em sua versão estática utilizando as abordagens inferenciais descritas no capítulo 4. Para isso, utilizaremos um conjunto de dados reais.

5.1 Conjunto de Dados

Os dados se referem aos rendimentos das Notas do Tesouro Nacional - Série B (NTN-B), calculados pela Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (Anbima), segundo própria metodologia¹, no período de janeiro de 2006 até dezembro de 2010, num total de 61 vencimentos, ou seja, o primeiro vencimento é após 15 dias ou 0,5 mês, o segundo vencimento é após 1 mês, o terceiro após 2 meses e assim sucessivamente durante todos os meses até o sexagésimo primeiro em dezembro de 2010. Os dados foram obtidos no dia 30/12/2010 e foi realizada uma interpolação para todos os períodos.

As NTN-B são títulos pós-fixados que tem seu valor nominal indexado mensalmente, desde a data-base, pela variação do IPCA.

¹Detalhes em http://www.andima.com.br/comites/arqs/com_anexo_6.pdf

A sigla IPCA significa Índice de Preços ao Consumidor Amplo, medido mensalmente pelo Sistema Nacional de Índices de Preços ao Consumidor do IBGE (Instituto Brasileiro de Geografia e Estatística). O índice mede a inflação dos preços de bens e serviços no varejo e foi escolhido porque é o índice mais relevante do ponto de vista da política monetária dado que o Conselho Monetário Nacional (CMN) estabelece a meta oficial de inflação em termos do IPCA.

O IPCA verifica as variações dos custos com os gastos das famílias com rendimento entre 1 (um) e 40 (quarenta) salários mínimos residentes nas regiões metropolitanas de Belém, Belo Horizonte, Brasília, Curitiba, Fortaleza, Porto Alegre, Recife, Rio de Janeiro, Salvador, São Paulo e município de Goiânia.

A pesquisa é realizada em estabelecimentos comerciais, prestadores de serviços, domicílios (para verificar valores de aluguel) e concessionárias de serviços públicos.

Índices de preços podem ser definidos como números que representam uma média ponderada dos preços de uma certa cesta de bens e serviços. Sua variação mede, portanto, a variação média (ponderada) dos preços dos produtos da cesta.

O IPCA considera nove grupos de produtos e serviços: alimentação e bebidas; artigos de residência; comunicação; despesas pessoais; educação; habitação; saúde e cuidados pessoais; transportes e vestuário. Esses nove grupos são subdivididos em itens, totalizando 465 subitens.

Os rendimentos das NTN-B estão dispostas na Tabela (5.1). Na 1ª coluna temos as maturidades em meses e na 2ª coluna o rendimento de cada maturidade definida como um percentual anual.

A Tabela 5.2 apresenta algumas estatísticas descritivas desse conjunto de dados.

Tabela 5.1: Rendimentos das NTN-B indexadas pelo IPCA

Maturidade	T. Juros	Maturidade	T. Juros	Maturidade	T. Juros
0,5	4,69	20	5,59	40	5,3
1	5,88	21	5,57	41	5,29
2	6,25	22	5,55	42	5,28
3	6,16	23	5,53	43	5,27
4	6,07	24	5,51	44	5,26
<hr/>					
Maturidade	T. Juros	Maturidade	T. Juros	Maturidade	T. Juros
5	6,01	25	5,49	45	5,26
6	5,97	26	5,47	46	5,25
7	5,94	27	5,45	47	5,24
8	5,91	28	5,44	48	5,24
9	5,89	29	5,42	49	5,23
10	5,86	30	5,41	50	5,23
11	5,83	31	5,39	51	5,22
12	5,8	32	5,38	52	5,21
13	5,77	33	5,37	53	5,21
14	5,75	34	5,36	54	5,2
15	5,72	35	5,34	55	5,2
16	5,69	36	5,33	56	5,2
17	5,67	37	5,32	57	5,19
18	5,64	38	5,31	58	5,19
19	5,62	39	5,3	59	5,18
				60	5,18

Tabela 5.2: Estatísticas descritivas dos rendimentos das NTN-B.

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
4,69%	5,25%	5,483%	5,39%	5,2%	5,69%	6,25%
<hr/>						
Medidas de Dispersão						
Desvio Padrão	Variância	Coef. Variação(%)				
0,3037	0,0922	5,5388				

Pelas estatísticas descritivas, nota-se que em média o rendimento para o período em estudo foi de 5,483% a.a. e que a distância interquartilica é de 0,44 ponto percentual, o que indica que temos uma grande concentração dos pontos em torno do

valor médio. Isso pode ser comprovado pelo coeficiente de variação que apresenta valor 5,539%. Por fim, temos que a variância dos dados também é baixa (0,0922), mais um indicativo da pouca dispersão dos dados.

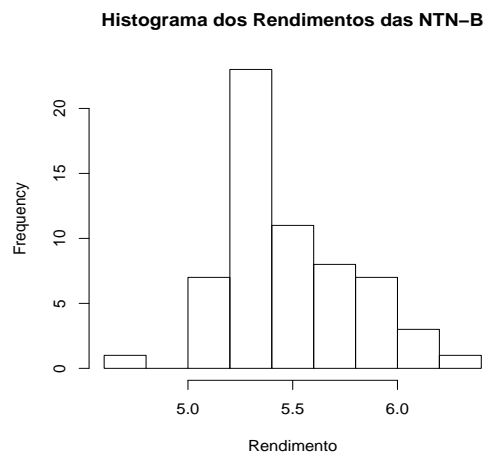


Figura 5.1: Histograma dos Rendimentos das NTN-B.

A Figura (5.1) mostra o histograma dos rendimentos das NTN-B pela qual verificamos que os dados estão distribuídos de forma assimétrica, o que é ratificado pelo coeficiente de assimetria de Pearson de 0,5068.

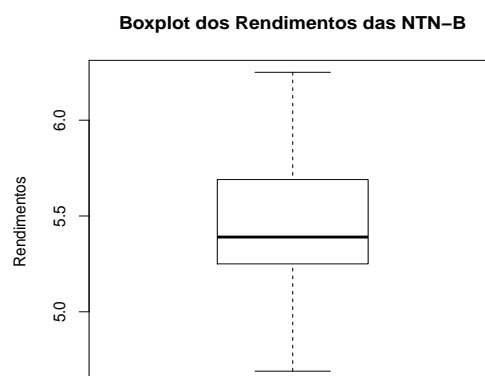


Figura 5.2: Boxplot

O boxplot (Figura (5.2)), não indica a presença de outliers ou pontos atípicos.

5.2 Estimação

Após a análise descritiva dos dados em estudo, iremos implementar as metodologias descritas no capítulo 4: inicialmente consideraremos o parâmetro de tempo (τ) fixo e num segundo momento esse parâmetro passará a ser estimado conjuntamente com os demais parâmetros.

Considerando o parâmetro de tempo fixo, adotaremos $\tau = 0,3985$. O motivo da escolha desse valor foi o fato de que quando ajustamos o conjunto de dados em estudo pelo *software R* no pacote *YieldCurve* verificamos que o modelo que apresentava o melhor resultado, teve o valor escolhido como estimativa para esse parâmetro.

5.2.1 Mínimos Quadrados Ordinários

Como vimos na Seção (4.1), o método dos Mínimos Quadrados Ordinários apresenta como melhor estimativa os valores dos parâmetros que apresentam a menor SQR e além disso, devemos verificar os dois casos válidos pelas condições KKT para encontrarmos as estimativas dos parâmetros.

Inicialmente utilizando o caso em que $\mu_1 < 0$ e $\mu_2 < 0$, temos que pelas condições KKT as estimativas para os parâmetros são:

$$\begin{cases} \hat{\beta}_0 = 0 \\ \hat{\beta}_1 = 0 \\ \hat{\beta}_2 = 34,24606 \end{cases}$$

Porém as estimativas dos parâmetros não podem assumir esses valores, pois as restrições $\mu_1 \leq 0$ e $\mu_2 \leq 0$ não são satisfeitas ($\hat{\mu}_1 = 920,1867$ e $\hat{\mu}_2 = 227,958$).

Portanto, devemos verificar o caso em que $\mu_1 = 0$ e $\mu_2 = 0$, assim obtemos as seguintes estimativas:

$$\begin{cases} \hat{\beta}_0 = 5,0600390 \\ \hat{\beta}_1 = -0,2394719 \\ \hat{\beta}_2 = 4,0276628 \end{cases}$$

Assumindo essas estimativas para os parâmetros, todas as condições KKT são satisfeitas e então conseguimos obter a curva que apresenta a menor SQR. Esperamos, como vimos na Seção (3.3), que a curva de taxas de juros apresente forma crescente para os títulos de curto prazo de vencimento ($\hat{\beta}_1 < 0$) e formato de “U invertido”, uma vez que a estimativa para o parâmetro β_2 é maior que zero. Dessa forma temos que a curva é dada por:

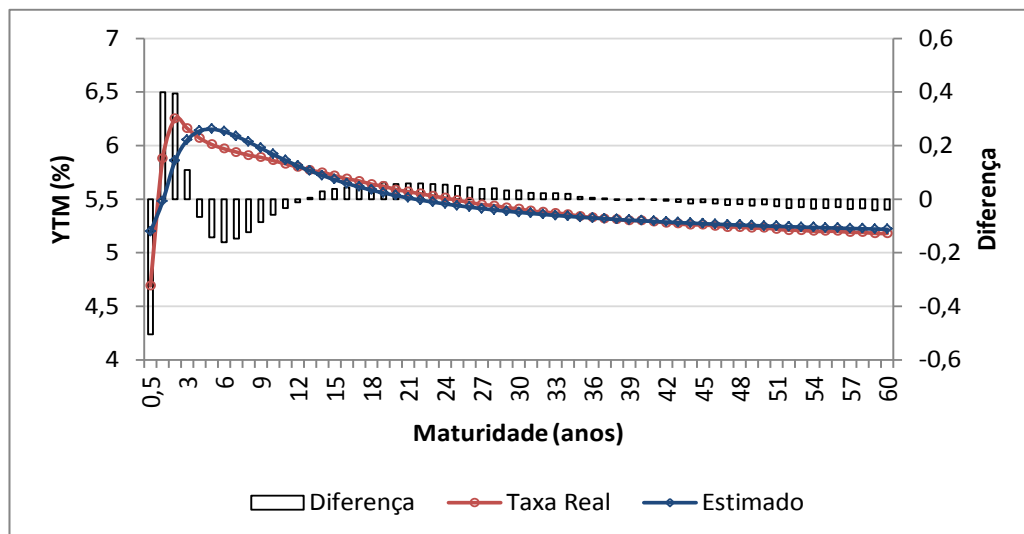


Figura 5.3: Gráfico da Curva de Taxas de Juros Ajustada (MQO)

A Figura (5.3) mostra o ajuste da curva de taxas de juros. Em seu eixo vertical principal temos os valores dos rendimentos para as diferentes maturidades (eixo horizontal) e no eixo vertical secundário temos a diferença entre a curva observada e a curva ajustada. Podemos observar que a estrutura da curva de taxas de juros tem o formato esperado, ou seja, apresenta forma crescente para os títulos de curto prazo de vencimento e formato de “U invertido”. Na Seção (2.2.2), vimos que por definição essa curva é definida como ondulada. Ainda podemos verificar que a curva ajustada e a observada estão praticamente sobrepostas indicando um bom ajuste ($SQR = 0,7371$).

5.2.2 Máxima Verossimilhança

Como vimos na Seção (4.2), o método da Máxima Verossimilhança necessita de uma hipótese sobre a distribuição das variáveis. Como anteriormente, assumiremos que os erros têm distribuição normal.

Além dos parâmetros de regressão (β_0 , β_1 e β_2) devemos estimar o parâmetro de dispersão (σ^2). Para estimarmos esses parâmetros novamente utilizaremos as condições KKT e avaliaremos os dois casos em que podemos obter os valores das estimativas.

No caso em que $\mu_1 < 0$ e $\mu_2 < 0$, novamente temos que as restrições $\mu_1 \leq 0$ e $\mu_2 \leq 0$ não são satisfeitas, pois assumem valores positivos.

Portanto, partiremos para o caso em que $\mu_1 = 0$ e $\mu_2 = 0$ encontrando as seguintes estimativas:

$$\left\{ \begin{array}{l} \hat{\beta}_0 = 5,060039 \\ \hat{\beta}_1 = -0,2394719 \\ \hat{\beta}_2 = 4,0276628 \\ \hat{\sigma}^2 = 0,01208349 \end{array} \right.$$

Como vimos no capítulo 4, as estimativas obtidas pelos métodos MQO e MV são as mesmas, portanto a Figura (5.3) também representa a estimativa obtida via MV.

5.2.3 Inferência Bayesiana Aplicada à ETTJ

Pela inferência bayesiana, conforme visto no capítulo 4, existem duas formas de estimação dos parâmetros do modelo. A primeira considerando o parâmetro de tempo (τ) fixo e a outra considerando como um parâmetro a ser estimado. Nas duas abordagens as restrições do modelo são incorporadas na definição das prioris.

- Estimação com τ fixo

Como na Seção (4.3), assumiremos a distribuição normal dos erros e as distribuições a priori. Definimos os hiperparâmetros da distribuição a priori de β_0 como $a=2$ e $b=20$, assim temos um Gama(2,20), para β_1 adotamos os valores $\mu_1=0$ e $\sigma_1^2=100$, assim temos uma distribuição Normal Truncada $(0, 100)$, para β_2 que tem distribuição Normal (μ_2, σ_2^2) , adotamos os valores $\mu_2=0$ e $\sigma_2^2=100$ e por fim temos o parâmetro σ^2 proporcional a $1/\sigma^2$.

O parâmetro β_0 não tem forma conhecida, como visto na Seção (4.3), por isso será necessário utilizar o algoritmo Métropolis-Hastings. Os demais parâmetros apresentam condicionais conhecidas e dessa forma, o amostrador de Gibbs é suficiente para obter as amostras a posteriori. Iremos realizar 1000 iterações, com *burn-in* de 1000 e saltos de 150, obtendo uma amostra final de 151000 valores gerados da posteriori.

No processo utilizando MCMC, devemos verificar se a cadeia obteve a convergência. Para essa verificação utilizamos o teste de Geweke, nele devemos

inicialmente descartar o *burn-in*, que é um aquecimento da cadeia até chegar mais próximo a uma região de alta probabilidade da estimativa real do parâmetro estar. Após este descarte o método consiste em utilizar partes da cadeia, normalmente 10% da parte inicial (A) e 50% da parte final (B) da cadeia. Definida as partes utilizadas, realizamos um teste de comparação de médias, em que a estatística de teste é dada por:

$$\frac{\bar{x}_A - \bar{x}_B}{\sqrt{\left(\frac{\hat{\sigma}_A^2}{n_A} - \frac{\hat{\sigma}_B^2}{n_B}\right)}} \quad (5.1)$$

onde \bar{x}_A e \bar{x}_B são as médias amostrais, $\hat{\sigma}_A^2$ e $\hat{\sigma}_B^2$ são as variâncias amostrais e n_A e n_B são os tamanhos das partes A e B da cadeia, respectivamente.

Em geral dizemos que a cadeia não convergiu se existir diferença entre as médias, ou seja, se a estatística teste (equação (5.1), estiver fora do intervalo de $[-2;2]$.

Para a amostra dos parâmetros obtida, o teste de Geweke apresenta os seguintes resultados para os parâmetros.

Tabela 5.3: Estatísticas do Teste de Geweke - τ fixo

β_0	β_1	β_2	σ^2
-0,652	0,8106	-0,4868	0,9783

Com o teste acima, verificamos a convergência da cadeia para os parâmetros envolvidos na análise, pois todos os valores se encontram dentro do intervalo de aceitação de convergência da cadeia.

Utilizando a Inferência Bayesiana obteremos uma amostra da posteriori e com base nessa amostra, obteremos as estatísticas descritivas (mínimo, média,

máximo, variância e intervalos HPD) e por algum critério específico, determinaremos a estimativa pontual para o parâmetro. Nesse trabalho, como estimativa pontual, foi adotado a média à posteriori.

Na Inferência Bayesiana quando adotamos como estimativa pontual para o parâmetro a média, significa que utilizamos como função de perda a quadrática, detalhes em Ehlers (2005).

Na Tabela (5.4) apresentamos as estatísticas descritivas da distribuição a posteriori para cada um dos parâmetros envolvidos na análise.

– Parâmetro β_0 :

Tabela 5.4: Estatísticas descritivas da amostra gerada para o parâmetro β_0 (τ fixo)

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
4,94	5,03	5,047	5,048	[5,04;5,06]	5,065	5,167
Medidas de Dispersão						
Desvio Padrão		Variância	Coef. Variação(%)			
0,0267		0,0007	0,5282			
Intervalos HPD						
90%		95%		99%		
(5,0051;5,0909)		(4,9966;5,1056)		(4,9667;5,1143)		

Pelas estatísticas descritivas, nota-se que para a amostra gerada o parâmetro β_0 temos o valor médio de 5,047, mediana de 5,048 e que a maior frequência do conjunto de dados (moda) está no intervalo entre 5,04 e 5,06. Além disso, podemos verificar que a variação da amostra gerada é pequena (0,007), o que indica a concentração da amostra em torno do valor médio. Pelos intervalos HPD, que nos mostram as regiões de

maiores densidades a posteriori com uma certa confiança, a amplitude é pequena nos intervalos, o que comprova a concentração da amostra em torno do valor médio.

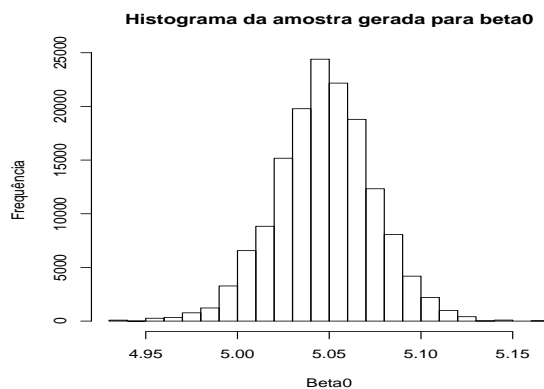


Figura 5.4: Histograma da amostra gerada para o parâmetro β_0 (τ fixo)

O histograma ilustrado na Figura (5.4) nos mostra como os valores gerados da posteriori para β_0 estão distribuídos. Como podemos observar, a distribuição se assemelha a uma densidade da distribuição normal, centrada no valor médio. Comprovamos a simetria da distribuição através do coeficiente de assimetria de Pearson, em que o valor da estatística é -0,0844, menor que o limite admitido para uma distribuição ser simétrica (0,15).

– Parâmetro β_1 :

Tabela 5.5: Estatísticas descritivas da amostra gerada para o parâmetro β_1 (τ fixo)

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
-0,706	-0,3129	-0,2440	-0,2434	[-0,45;-0,2]	-0,1744	0,2483

Medidas de Dispersão		
Desvio Padrão	Variância	Coef. Variação(%)
0,1032	0,0107	-42,3146

Intervalos HPD		
90%	95%	99%
(-0,4055;-0,0857)	(-0,4288;-0,0569)	(-0,4897;0,0112)

Destacamos a média de $-0,2440$, que é adotada como estimativa pontual para o parâmetro. Apesar da estimativa ser negativa, as restrições do modelo continuam validas, pois $\hat{\beta}_0=5.047$, como visto pela Tabela (5.4), dessa forma $\beta_0 + \beta_1 \geq 0$. A variância baixa indica a concentração da amostra em torno do valor médio. Para os intervalos HPD temos que com uma probabilidade de cobertura de 99%, o intervalo $[-0,4897;0,0112]$ é o de menor comprimento, para 95% o intervalo de menor comprimento é $[-0,4288;-0,0569]$ e para 90%, temos que o intervalo de menor comprimento é dado por $[-0,4055;-0,0857]$.

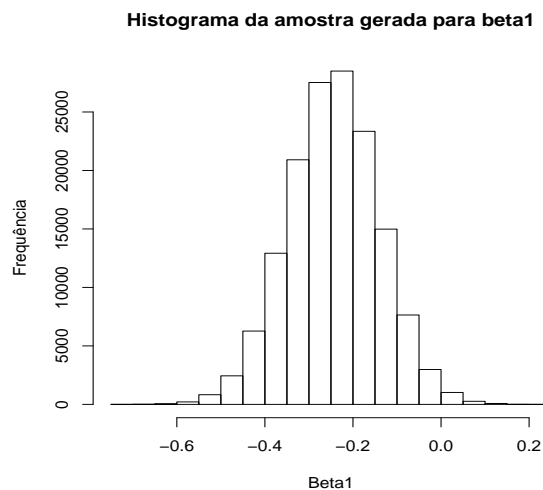


Figura 5.5: Histograma da amostra gerada para o parâmetro β_1 (τ fixo)

Calculando o coeficiente de assimetria de Pearson, obtemos o valor de -0.0266 , o que nos comprova que a distribuição é simétrica. Podemos dizer que o formato da distribuição se assemelha a densidade da distribuição da Normal.

– Parâmetro β_2 :

Tabela 5.6: Estatísticas descritivas da amostra gerada para o parâmetro β_2 (τ fixo)

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
2,947	3,94	4,109	4,107	[4;4,4]	4,277	5,408
Medidas de Dispersão						
Desvio Padrão	Variância	Coef. Variação(%)				
0,2538	0,0644	6,1772				
Intervalos HPD						
90%		95%		99%		
(3,7085;4,4863)		(3,6160;4,5802)		(3,3922;4,7048)		

Notamos que as medidas de média, mediana e moda são aproximadamente iguais, apresentando diferença apenas na terceira casa decimal. Esse fato é característico quando temos a densidade da distribuição normal (média=mediana=moda). Já as medidas de dispersão nos mostram que a amostra se encontra em torno do valor médio. Os intervalos HPD, com suas respectivas probabilidades de cobertura, apresentam a densidade para qualquer ponto pertencente ao intervalo sempre maior que a densidade para qualquer ponto não pertencente ao intervalo.

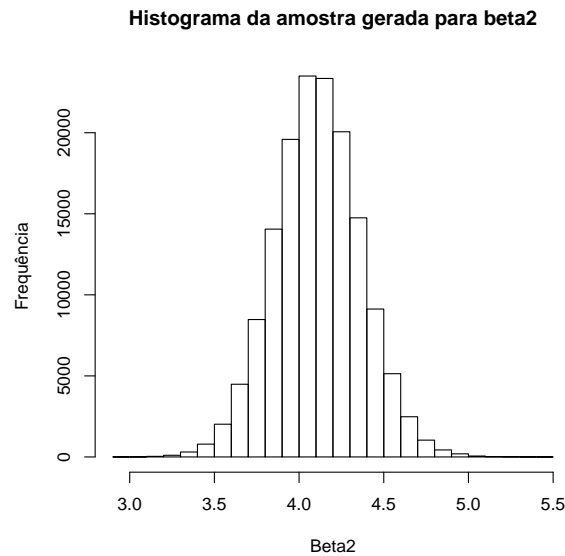


Figura 5.6: Histograma da amostra gerada para o parâmetro β_2 (τ fixo)

Como esperado o histograma é semelhante à densidade da distribuição normal, centrada na média=mediana=moda. A simetria da distribuição é comprovada através do coeficiente de assimetria de Pearson, igual a 0,0517.

– Parâmetro σ^2 :

Tabela 5.7: Estatísticas descritivas da amostra gerada para o parâmetro σ^2 (τ fixo)

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
0,0063	0,0114	0,0132	0,0129	[0,012;0,014]	0,01467	0,03583

Medidas de Dispersão		
Desvio Padrão	Variância	Coef. Variação(%)
0,0026	0,000006	19,3598

Intervalos HPD		
90%	95%	99%
(0,0091;0,0165)	(0,0091;0,0180)	(0,0086;0,0208)

Por fim, para o parâmetro de dispersão, podemos notar que ele apresenta valores aproximadamente zero, com seus valores bem concentrados em torno do valor médio de 0,0132. Isso pode ser confirmado pelos valores de variância, desvio padrão e pela diferença interquartílica. Nos intervalos, temos uma amplitude pequena, comprovando o que foi observado através das estatísticas descritivas, a concentração da amostra em torno do valor médio.

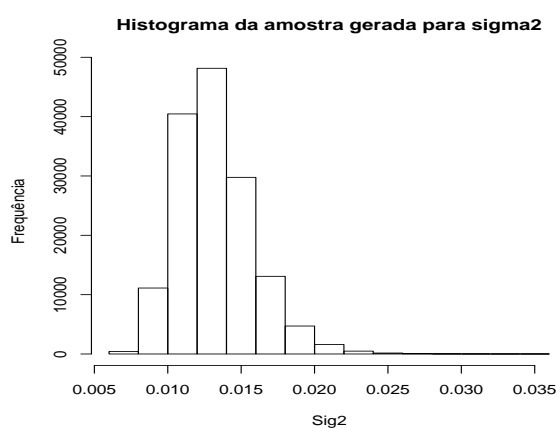


Figura 5.7: Histograma da amostra gerada para o parâmetro σ^2 (τ fixo)

Como estamos trabalhando com um parâmetro de dispersão, esperamos que ele apresente uma assimetria a direita. Pelo histograma (Figura (5.7)) essa assimetria pode ser claramente observada. Comprovamos a assimetria através do valor de 0,8241, calculado para o coeficiente de assimetria de Pearson, que está classificado na faixa de uma distribuição assimétrica moderada. Por se tratar de um parâmetro de variância, sabemos que ele é positivo e observando o formato do histograma, podemos dizer que o formato se aproxima de uma distribuição Gama.

Como já mencionado anteriormente, neste trabalho adotamos como estimativa pontual para os parâmetros do modelo a média a posteriori. Dessa forma, obtida as estimativas dos parâmetros, podemos representar graficamente a curva de taxas de juros e também comparar esses resultados com os resultados obtidos pelos métodos frequentistas.

Tabela 5.8: Comparação das estimativas dos parâmetros pelos métodos apresentados.

Mínimos Quadrados	Máxima Verossimilhança	Inferência Bayesiana
$\hat{\beta}_0 = 5,0600390$	$\hat{\beta}_0 = 5,0600390$	$\hat{\beta}_0 = 5,047$
$\hat{\beta}_1 = -0,2394719$	$\hat{\beta}_1 = -0,2394719$	$\hat{\beta}_1 = -0,2440$
$\hat{\beta}_2 = 4,0276628$	$\hat{\beta}_2 = 4,0276628$	$\hat{\beta}_2 = 4,109$
	$\hat{\sigma}^2 = 0,01208349$	$\hat{\sigma}^2 = 0,0132$

Pela tabela, nota-se que todas as estimativas estão próximas, apresentando diferença apenas nas casas decimais, o que confirma que o método utilizando inferência bayesiana se aproxima dos métodos frequentistas normalmente utilizados. Essa igualdade também pode ser verificada através da SQR, em que por todos os métodos o valor calculado foi de aproximadamente 0,74. Com base no ajuste desse conjunto de dados e pelas simulações realizadas no capítulo 4, podemos dizer que a abordagem bayesiana, utilizando as prioris definidas no trabalho, é uma outra e boa forma de obtermos as estimativas dos parâmetros do modelo. Comprovamos a qualidade do ajuste através da representação gráfica.

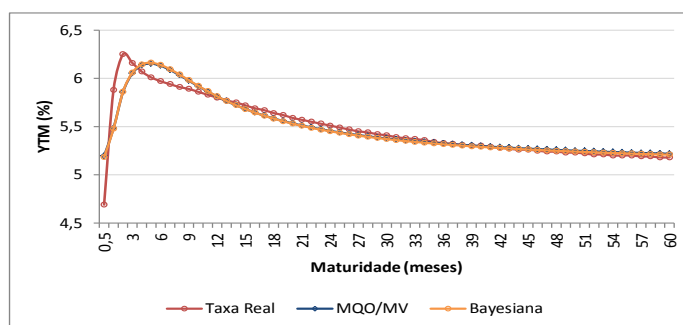


Figura 5.8: Curva de taxa de juros - Dados reais - τ fixo

Pelo gráfico, podemos ver que as curvas construídas com as estimativas dos parâmetros encontradas pelos métodos MQO, MV e Inferência Bayesiana se encontram praticamente sobrepostas sobre a curva real indicando que os métodos são boas formas de obtermos, por exemplo, qual o rendimento para uma maturidade fora das praticadas no mercado para esse ativo.

- Estimação com τ estimado

Os métodos apresentados até o momento nesse capítulo consideraram o parâmetro de tempo τ fixo, porém como vimos no capítulo três, temos outras formas de definirmos esse parâmetro e uma delas é estimá-lo juntamente com os demais parâmetros. Para realizar essa estimação utilizaremos os resultados apresentados na Seção 4.4.

Temos novamente que utilizar um recurso MCMC para obter amostras a posteriori para os parâmetros. Para isso foi utilizado para os parâmetros β_0 e τ o algoritmo de Metropolis-Hastings dado que não temos uma condicional conhecida e para os outros parâmetros envolvidos na análise, por apresentar

condicionais conhecidas, foi utilizado o amostrador de Gibbs.

Devemos definir os hiperparâmetros das prioris apresentadas na Seção 4.4.

- $\beta_0 \sim \text{Gama}(2,20)$.
- $\beta_1 \sim \text{Normal}(-10,100)$.
- $\beta_2 \sim \text{Normal}(0,100)$.
- $\tau \sim \text{Gama}(1,2)$.
- σ^2 a priori proporcional a $1/\sigma^2$.

Realizamos 3000 iterações, com *burn-in* de 1000 e saltos de 10, obtendo uma amostra final de 31000 valores gerados da posteriori. Como estamos utilizando um procedimento MCMC, devemos verificar a convergência da cadeia, que é feito pelo teste de Geweke, descrito pela equação (5.1). Os resultados são apresentados na Tabela (5.13).

Tabela 5.9: Estatísticas do Teste de Geweke - τ estimado

β_0	β_1	β_2	σ^2	τ
-1,318	-1,452	1,54	-1,516	-0,399

Verificamos que a convergência da cadeia foi obtida. Com a convergência obtida, iremos agora realizar um estudo da amostra gerada da posteriori para cada parâmetro. Para não tornar repetitivo as tabelas e gráficos vistas no método considerando τ fixo, iremos apresentar apenas o valor da média a posteriori, que é utilizado neste trabalho como estimativa pontual, a variância como medida de dispersão dos valores da amostra e o intervalo HPD para os parâmetros β_0 , β_1 , β_2 e σ^2 .

Tabela 5.10: Estatísticas da amostra gerada para os parâmetros β_0 , β_1 , β_2 e σ^2 - τ estimado

		β_0	β_1	β_2	σ^2
	Média	5,152	-1,032	5,314	0,0109
	Variância	0,000001	0,0161	0,0536	0,000004
Int. HPD	90%	[5,151;5,154]	[-1,250;-0,829]	[4,94;5,697]	[0,008;0,014]
	95%	[5,150;5,154]	[-1,285;-0,787]	[4,862;5,759]	[0,007;0,015]
	99%	[5,149;5,155]	[-1,343;-0,698]	[4,749;5,914]	[0,006;0,017]

Na Tabela (5.10) temos os valores da média a posteriori, em que nesse trabalho são considerados a estimativa pontual para os parâmetros, além do valor das variâncias das amostras, que observamos que para todos os parâmetros apresentam valor baixo, indicando assim a concentração das amostras em torno do valor médio. Pelos intervalos HPD podemos verificar com uma certa confiança as regiões de maiores densidades a posteriori, onde temos que os intervalos apresentam amplitudes pequenas, comprovando a concentração.

Para o parâmetro de tempo (τ), temos que as estatísticas obtidas através da amostra gerada da posteriori é dada por:

- Parâmetro τ :

Tabela 5.11: Estatísticas descritivas da amostra gerada para o parâmetro τ (estimado)

Medidas de Posição						
Mínimo	1ºQuartil	Média	Mediana	Moda	3ºQuartil	Máximo
0,5976	0,6327	0,6389	0,6389	[0,635;0,645]	0,6453	0,6737

Medidas de Dispersão		
Desvio Padrão	Variância	Coef. Variação(%)
0,0095	0,00008	1,4810

Intervalos HPD		
90%	95%	99%
(0,6232;0,6539)	(0,6204;0,6566)	(0,6155;0,6633)

Por fim, para o parâmetro de tempo que agora é uma variável que estimamos, temos que a média, mediana e moda são iguais a 0,6389. Temos ainda que a amplitude de valores da amostra a posteriori é de 0,0761 (máximo=0,6737 - mínimo=0,5976), o que indica que os dados estão concentrados em torno do valor médio. Esse fato é comprovado pelas medidas de dispersão, em que a variância é 0,00008 e o coeficiente de variação é de 1,4810. Nos intervalos temos uma amplitude pequena, comprovando o fato da concentração da amostra em torno do valor médio.

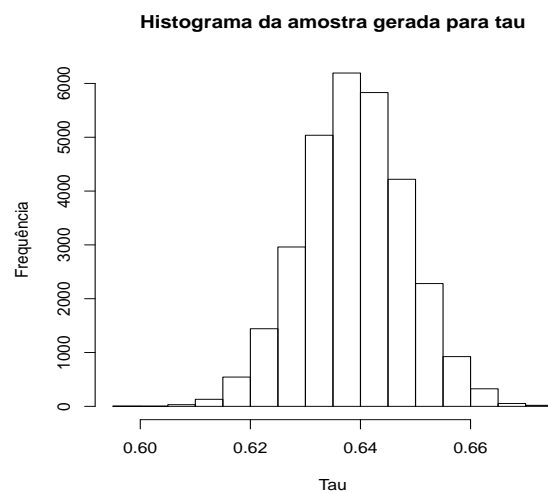


Figura 5.9: Histograma da amostra gerada para o parâmetro τ (τ estimado)

No histograma podemos verificar como a amostra está distribuída. Podemos

notar que apresenta o formato da densidade de uma distribuição normal.

Com as estatísticas obtidas da amostra gerada da posteriori para cada parâmetro, utilizamos para construção da curva de taxa de juros a média à posteriori dada pelas Tabelas (5.10) e (5.11) que são:

Tabela 5.12: Estimativas dos parâmetros - Inferência Bayesiana - τ estimado

Inferência Bayesiana
$\hat{\beta}_0 = 5,152$
$\hat{\beta}_1 = -1,032$
$\hat{\beta}_2 = 5,314$
$\hat{\sigma}^2 = 0,0109$
$\hat{\tau} = 0,6389$

Com as estimativas pontuais, conseguimos construir a curva de taxa de juros, mostrada na figura a seguir:

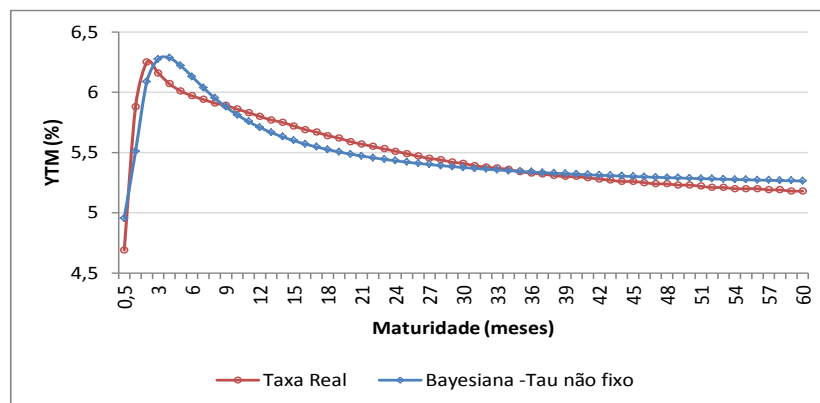


Figura 5.10: Curva de taxa de juros - Dados reais - τ estimado

Podemos ver novamente que a curva de taxa de juros construída com base nas estimativas encontradas, utilizando a inferência bayesiana, está praticamente sobreposta a curva de taxa de juros real, o que indica um bom ajuste da curva por esse método.

Comparando esse ajuste com os demais ajustes apresentados nesse trabalho, vemos que estimar o parâmetro de tempo juntamente com os demais parâmetros apresenta resultados melhores para esse conjunto de dados, considerando a SQR como critério, pois a SQR da curva ajustada por esse método apresenta valor de 0,62, enquanto que pelos outros métodos foi de aproximadamente 0,74.

Outra vantagem em estimarmos o parâmetro de tempo juntamente com os demais fatores é que quando utilizamos ele como um valor fixo, a única informação que temos para determinar seu valor é o fato que, quando ele assume um valor pequeno a curva se ajusta melhor a títulos de curto prazo e quando assume valor grande a curva se ajusta melhor para títulos de longo prazo. Contudo, mesmo com essa informação é adotado um valor arbitrário para ele, resultando em ajustes menos precisos que quando estimamos ele juntamente com os demais parâmetros.

Capítulo 6

Conclusão

Neste trabalho foi apresentado uma análise estatística do modelo de Nelson e Siegel. Este é o modelo de maior utilização no Brasil e no mundo quando necessitamos modelar a curva de taxas de juros, pois como foi mostrado, é um modelo de fácil implementação, parcimonioso e que se ajusta aos diversos formatos que as curvas de taxas de juros possuem.

Por ser um modelo bastante utilizado, a estimação dos seus parâmetros deve ser realizada de modo preciso, ou seja, deve representar de forma fiel a curva real pois esta curva será utilizada como uma ferramenta de extrema importância em decisões financeiras.

Por ser de extrema importância a representação fiel da curva, nosso objetivo no estudo foi de apresentar novas metodologias para estimar os parâmetros do modelo levando em conta suas restrições. Dentro destas novas formas, foram utilizados os métodos frequentistas, tradicionalmente empregados em modelos lineares (MV e MQO) e além destes métodos tradicionais foi introduzido a Inferência Bayesiana, na qual consideramos o parâmetro de tempo (τ) fixo e estimado conjuntamente aos demais parâmetros.

Este objetivo de estudo, que consiste em mostrar novas metodologias para

obter as estimativas dos parâmetros, surgiu com o ajuste do conjunto de dados apresentado no trabalho em que o programa existente no *software R* não respeitava as restrições do modelo e além disso o ajuste não era fiel a curva real.

Inicialmente utilizamos as metodologias clássicas (MV e MQO) para buscar as estimativas dos parâmetros. Nesta forma, consideramos o parâmetro de tempo (τ) fixo e como tínhamos restrições de desigualdade, foi utilizado como suporte para obtermos as estimativas o método KKT. Pelos estudos apresentados, notamos que o método de MQO e MV são eficazes, sendo necessário apenas os ajustes para o caso em que $\mu_1 = 0$ e $\mu_2 = 0$ e para o caso em que $\mu_1 < 0$ e $\mu_2 < 0$. Desta forma, pelas restrições do método, apenas um caso é aceito, e assim as melhores estimativas para os parâmetros são encontradas.

Na sequência, foi utilizada a Inferência Bayesiana, na qual através da definição das prioris incorporamos as restrições do modelo. Neste trabalho definimos as prioris, contudo outras prioris que respeitem as restrições poderiam ser utilizadas, como por exemplo o uso da priori log-normal para o parâmetro de tempo ou o uso de uma gaussiana para o parâmetro β_0 . Neste estudo, os resultados encontrados foram satisfatórios com essas prioris, porém quando utilizamos a versão dinâmica do modelo, os resultados podem não ser satisfatórios com essas prioris.

Novamente iniciamos a utilização da Inferência Bayesiana para o caso em que o parâmetro de tempo (τ) é fixo. E foi verificado tanto pelos estudos de simulação como também pelo ajuste do conjunto de dados real, que o método encontra com precisão as estimativas dos parâmetros.

Por fim, consideramos o parâmetro de tempo (τ) estimado juntamente com os demais parâmetros e assim como no caso em que era considerado fixo, os resultados obtidos foram satisfatórios. Porém, como visto anteriormente, utilizando como critério a SQR e considerando o método com τ estimado, os resultados foram melhores no ajuste do conjunto de dados reais, pois encontramos SQR=0,62

enquanto que pelos outros métodos foi de aproximadamente 0,74.

Uma vez que os resultados foram satisfatórios quando τ foi considerado outro parâmetro a ser estimado, é recomendado a estimação conjunta dos parâmetros, pois quando utilizamos o mesmo fixo a única informação que temos para determinar seu valor é o fato de que, quando ele assume um valor pequeno, a curva se ajusta melhor a títulos de curto prazo e quando assume valor grande, a curva se ajusta melhor para títulos de longo prazo. Entretanto, mesmo com essa informação é adotado um valor arbitrário para o mesmo, resultando em ajustes menos precisos que quando estimado juntamente com os demais parâmetros.

A utilização do parâmetro de tempo (τ) fixo é recomendada apenas em casos específicos quando não estamos interessados na curva como um todo, apenas em títulos de curto prazo ou de longo prazo. Desta forma, atribuindo um valor fixo e arbitrário (pequeno ou grande, de acordo com a preferência do agente de mercado) a curva resultante neste período pode ser mais precisa do que o ajuste da curva total.

Por todas essas considerações neste trabalho, visamos apresentar novas metodologias para obtermos as estimativas dos parâmetros do modelo, que levam em conta as restrições. Pelo fato de incorporar as restrições, os resultados são mais confiáveis e assim temos uma forma mais segura de representar a curva de taxas de juros.

Apêndice A

Distribuições Utilizadas

- Distribuição Normal:

Dizemos que X apresenta distribuição normal com média μ e variância σ^2 , denotando-se $X \sim N(\mu, \sigma^2)$, se sua função de densidade for da seguinte forma:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

com $-\infty < x < \infty$ e $\sigma^2 > 0$.

- Distribuição Gama:

X tem distribuição Gama com parâmetros a e b , se sua densidade é dada por:

$$\frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx),$$

$x > 0$.

E temos que a média dessa distribuição é $E(X) = a/b$, e a variância é $Var(X) = a/b^2$.

- Distribuição Gama Inversa:

X tem distribuição Gama Inversa com parâmetros a e b , se sua densidade é dada por:

$$\frac{b^a}{\Gamma(a)} x^{-a+1} \exp(-b/x),$$

$x > 0$.

E temos que a média dessa distribuição é $E(X) = b/(a-1)$, e a variância é $Var(X) = \frac{b^2}{(a-1)^2(a-2)}$.

- Distribuição Normal Truncada:

Dizemos que X apresenta distribuição normal truncada em x' , com média μ e variância σ^2 , denotando-se $X \sim NT(\mu, \sigma^2)$, se sua função de densidade for da seguinte forma:

$$\left[\frac{1}{1 - \Phi(x')} \right] \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right),$$

sendo Φ a função distribuição acumulada da distribuição normal até o ponto x' .

Referências Bibliográficas

- [1] BAYES, T. *An essay towards solving in the doctrine of chances*, PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY LONDON 53, 1763, 370:418.
- [2] BIS, *Zero-coupon yield curves: Technical documentation*, BANK FOR INTERNATIONAL SETTLEMENTS PAPERS, 2005.
- [3] CARRIERE, J. F.; *Long-term yield rates for actuarial valuations*, ACTUARIAL FILE LIBRARY, 1998.
- [4] COSTA NETO, A. S. *Cálculo II*, FACULDADE DE TECNOLOGIA E CIÊNCIAS - ENSINO A DISTÂNCIA, 2006.
- [5] COURANT, E.; HILBERT, D., *Methods of Mathematical Physics*, 1953, 1:93-97.
- [6] CASELLA, G.; BERGER, R. L., *Statistical Inference*, 2001, 2^A EDIÇÃO.
- [7] COX, J. C.; INGERSOLL, J. E.; ROSS, S. A., *A theory of the term structure of interest rates*, ECONOMETRICA, 1985, VOL. 53, PP. 385:407.
- [8] CSAJBOK, A. *Zero-coupon yield curve estimation from a central bank perspective*, NATIONAL BANK OF HUNGARY WORKING PAPER, 1998.
- [9] CULBERTSON, J. M., *The Term Structure of Interest Rates*, THE QUARTERLY JOURNAL OF ECONOMICS, VOL.71, NO.4, 11/1957, PP. 485:517.

- [10] DIEBOLD, F. X.; LI, C., *Forecasting the term structure of government Bond yields*, JOURNAL OF ECONOMETRICS 130, 2006, PP. 337:364.
- [11] DOBBIE, G. M; WILKIE, A. D., *The Financial Times Actuaries Fixed Interest Indices*, TRANSACTION OF THE FACULTY OF ACTUARIES, 1979, VOL. 36, PP. 203:213..
- [12] EHLERS, R.S. *Introdução à Inferência Bayesiana*, DEPARTAMENTO DE ESTATÍSTICA, UFPR, DISPONÍVEL EM [HTTP://LEG.EST.UFPR.BR/ EHLERS/NOTAS/BAYES2006](http://leg.est.ufpr.br/~ehlers/notas/bayes2006), 2005.
- [13] FABOZZI, F. J.; FOCARDI, S. M., *The Mathematics of Financial Modeling and Investment Management*, FRANK J. FABOZZI SERIES, 2004.
- [14] FISHER, I., *Appreciation and Interest*, THE AMERICAN ECONOMIC ASSOCIATION, VOL.XI, No.4, 1896, PP. 331:442.
- [15] GAMERMAN, D.; *Markov Chain Monte Carlo: stochastic simulation for bayesian inference*, CHAPMAN & HALL, 1997.
- [16] GRAYBILL, F. A.; *Theory and Application of the Linear Model*, DUXBURY, 2000.
- [17] GOURIEROUX, C.; MONFORT, A., *Statistics and econometric models*, NEW YORK: CAMBRIDGE, 1995. 2 v.
- [18] GUEDES, J.; *Uma aplicação da abordagem de Nelson e Siegel (1987) reformulada por Diebold e Li (2006) e comparação com a Análise de Componentes Principais*, INSTITUTO DE GESTÃO DA TESOURARIA E DO CRÉDITO PÚBLICO, NÚCLEO DE CONTROLO FINANCEIRO, 2008.
- [19] HICKS, J. R., *Value and Capital an inquiry into some fundamental principles of economic theory*, CLARENDON PRESS IN OXFORD, 1946.

- [20] IOANNIDES, M. *A comparison of yield curve estimation techniques using UK data*, JOURNAL OF BANKING AND FINANCE 27, 2003, 1:26.
- [21] KARUSH, W., *Minima of Functions of Several Variables with Inequalities as Side Constraints*, M.Sc. DISSERTATION. DEPT. OF MATHEMATICS, UNIV. OF CHICAGO, CHICAGO, ILLINOIS, 1939.
- [22] KUHN, H. W.; TUCKER, A. W., *Nonlinear programming*, PROCEEDINGS OF 2ND BERKELEY SYMPOSIUM: 481:492, BERKELEY: UNIVERSITY OF CALIFORNIA PRESS, 1951.
- [23] LAURINI, M. P., HOTTA, L. K.; *Extensões Bayesianas do Modelo de Estrutura a Termo de Diebold-Li*, 2007.
- [24] LAURINI, M. P., NETO, A. D. W.; *Arbitragem na Estrutura a Termo das Taxas de Juros: Uma Abordagem Bayesiana*, INSPER WORKING PAPER, 2010.
- [25] LITTERMAN, R., SCHEINKMAN, J.; *Common Factors Affecting Bond Returns*, JOURNAL OF FIXED INCOME, VOL. 1, NO. 1, PP. 54:61, JUN 1991.
- [26] MANOUSOPOULOS, P.; MICHALOPOULOS, M., *Comparison of non-linear optimization algorithms for yield curve estimation*, EUROPEAN JOURNAL OF OPERATIONAL RESEARCH, 2009, 594:602.
- [27] McCULLOCH, J. H., *Measuring the Term Structure of Interest Rates*, THE JOURNAL OF BUSINESS, UNIVERSITY OF CHICAGO PRESS, VOL. 44(1), PP. 19:31, JAN 1971.
- [28] McCULLOCH, J. H., *An Estimate of the Liquidity Premium*, JOURNAL OF POLITICAL ECONOMY, UNIVERSITY OF CHICAGO PRESS, VOL. 83(1), PP. 95:119, FEV 1975.

- [29] MOOD, A.M.; GRAYBILL, F.A.; BOES, D.C.; *Introduction to the theory of statistics*, MCGRAW-HILL, 1974.
- [30] NELSON, C. R.; SIEGEL, A. F., *Parsimonious Modeling of Yield Curves*, *The Journal of Business*, VOLUME 60, NÚMERO 4 (OUTUBRO, 1987), PP. 473:489.
- [31] PAULINO, C.D, TURKMAN, M.A.A. E MURTEIRA, B; *Estatística Bayesiana*, FUNDAÇÃO CALOUSTE GULBENKIAN, LISBOA, 2003.
- [32] ROBERT, C. P.; CASELLA, G., *Monte Carlo Statistical Methods*, SPRINGER, 1999.
- [33] STANDER, Y. S., *Yield Curve Modeling*, PALGRAVE MACMILLAN, 2005.
- [34] SVENSSON, L.E.O., *Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994*, PAPERS 579, STOCKHOLM - INTERNATIONAL ECONOMIC STUDIES, 1994.
- [35] VASICEK, O. A., *An equilibrium characterisation of the term structure*, JOURNAL OF FINANCIAL ECONOMICS, 1977, VOL. 5, PP. 177:188.
- [36] VASICEK, O. A.; FONG, H. G., *Term Structure Modeling Using Exponential Splines*, JOURNAL OF FINANCIAL, 1982, VOL. 37, PP. 339:348.
- [37] WARSH, K. M., *Financial Markets and the Federal Reserve*, NEW YORK STOCK EXCHANGE, NEW YORK, NY, 21/11/2006.