

Estimação dos Parâmetros da Estrutura a Termo das Taxas de Juros Utilizando um Algoritmo de Evolução Diferencial

Resumo

A estimação da estrutura a termo das taxas de juros é um dos tópicos de pesquisa mais importante em macroeconomia e finanças, uma vez que a curva de juros é instrumento essencial na precificação de fluxos de caixa financeiros, permite a avaliação da efetividade de decisões de políticas monetárias, e fornece expectativas do mercado quanto ao futuro das taxas de juros, inflação e atividade econômica. Este trabalho propõe a utilização de um algoritmo de Evolução Diferencial para estimar os parâmetros das curvas de Nelson e Siegel (1987) e Svensson (1994), modelos parcimoniosos que descrevem a curva de juros zero-cupom, com base em dados das Letras do Tesouro Nacional, para o período de Janeiro de 2007 a Dezembro de 2009. Os resultados mostram que o algoritmo evolutivo proposto fornece estimativas mais precisas para as curvas de juros em termos de significância estatística, superando a tradicional técnica de estimação de mínimos quadrados não-lineares. Além disso, foi verificada uma menor variabilidade dos parâmetros estimados por meio da metodologia de evolução diferencial, o que confirma sua adequação aos princípios financeiros por trás da construção da estrutura a termo.

Palavras-Chave: Curva de Juros, Evolução Diferencial, Modelos Nelson-Siegel, Computação Evolutiva, Otimização.

Abstract

The term structure of interest rates estimation is one of the most important research topics in macroeconomics and finance, since the yield curve is an essential tool for pricing financial cash flows, allows the effectiveness evaluation of monetary policy decisions, and provides market expectations about the future of interest rates, inflation and economic activity. This work suggests the use of an Differential Evolution algorithm to estimate the parameters of the Nelson and Siegel (1987) and Svensson (1994) curves, parsimonious models that describe zero-coupon yield curve, based on Brazilian National Treasury Bills, for the period from January 2007 to December 2009. The results show that the evolutionary computation algorithm proposed provides more accurate estimates of the yield curves in statistical terms, outperforming the traditional nonlinear least squares technique. Parameters estimate by using differential evolutionary the methodology also exhibited lower variability, which confirms its adequacy to the financial principles behind term structure construction.

Keywords: Interest Rate Curve, Differential Evolution, Nelson-Siegel Models, Evolutionary Computation, Optimization.

Área ANPEC: Área 7 - Microeconomia, Métodos Quantitativos e Finanças

Classificação JEL: G12, E43, C61.

1 Introdução

A curva de rendimentos descreve a evolução temporal da estrutura a termo das taxas de juros (ETTJ) e é uma importante ferramenta utilizada em decisões financeiras e macroeconômicas. Em particular, a curva de juros é empregada na precificação de títulos financeiros, gerenciamento de riscos, decisões financeiras corporativas, e construção de portfólios. Além disso, desempenha um papel fundamental na implementação e avaliação de políticas monetárias, assim como é capaz de indicar as expectativas dos agentes de mercado quanto ao comportamento futuro das taxas de juros, inflação e atividade econômica.

A relação funcional entre as taxas nominais de juros de títulos zero-cupom e maturidade, de acordo com a hipótese de não-arbitragem da estrutura a termo, deve apresentar continuidade e suavidade¹. Sobretudo, a ausência de títulos públicos para todas as maturidades implica que, na prática, a estrutura a termo não é observada, devendo ser estimada utilizando um número limitado de taxas à vista diretamente observadas dos títulos zero-cupom e os respectivos preços desses títulos.

Métodos baseados em fórmulas funcionais, isto é, expressões matemáticas capazes de descrever toda estrutura das taxas de juros, que sejam parcimoniosos nos parâmetros e fáceis de implementar vêm sendo desenvolvidos por pesquisadores e amplamente utilizados por bancos centrais. Por exemplo, o modelo desenvolvido por (Nelson e Siegel 1987) e sua versão estendida, proposta por (Svensson 1994) (funções Nelson-Siegel-Svensson ou NSS), são atualmente utilizados por nove dos treze bancos centrais que fornecem suas estimativas da estrutura a termo para o *Bank of International Settlements* (BIS 2005)².

A estimação da curva de juros, de acordo com as funções de NSS, é essencialmente um problema de otimização não-linear, o que implica no risco de falsa convergência, i.e., possibilidade de que os parâmetros obtidos possam corresponder a um ponto de ótimo local e não de ótimo global, devido à não-linearidade nos parâmetros, resultando em curvas menos precisas. Além disso, observa-se uma sensibilidade significativa dos parâmetros às condições iniciais utilizadas na implementação dos algoritmos tradicionais de otimização não-linear (máxima verossimilhança e mínimos quadrados não-lineares). A principal consequência do problema de convergência é que, de acordo com as condições iniciais, as estimativas dos parâmetros exibem alta variabilidade.

Para tornar as funções de Nelson-Siegel-Svensson lineares nos parâmetros, muitos trabalhos fixam, arbitrariamente, os parâmetros que descrevem a velocidade para se atingir a taxa de longo prazo, permitindo a aplicação da técnica de mínimos quadrados ordinários (Diebold e Li 2006, Diebold, Li e Yue 2008). Entretanto, conforme apontado por (Bolder e Strélski 1999), a estimação da curva de juros tende a ser muito instável ao se avaliar uma função custo de alta não-linearidade e, ao impor a linearização ao modelo, pode-se gerar curvas de juros ainda mais instáveis, principalmente em países emergentes, como o Brasil; além de fornecer estimativas que não refletem a condição de mercado avaliada.

A literatura, em geral, discute a estimação de diferentes modelos funcionais para a curva de juros, mas não o método de otimização empregado. Por exemplo, (Ioannides 2003) avaliou diversos modelos que descrevem a ETTJ utilizando o algoritmo BFGS³, enquanto que (Csajbok 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 juros, baseados em dados reais de mercado. Os autores mostraram que alguns algoritmos são mais efetivos dependendo da natureza do problema e, além disso, a precisão aumenta quando se considera todos os parâmetros variantes no processo de otimização.

Adotando uma nova abordagem, baseados em técnicas de Algoritmos Evolutivos (AEs), (Gimeno e Nave 2009) propuseram a utilização de algoritmos genéticos para a estimação dos parâmetros das funções de Nelson-Siegel-Svensson no mercado de títulos públicos espanhol. Segundo os autores, essa abordagem permite a superação da limitação associada ao risco de falsa convergência, observada nas tradicionais técnicas de programação não-linear. Os resultados mostraram que o algoritmo genético proposto permitiu a construção de uma curva de juros mais precisa e, além disso, gerou uma série menos volátil

¹ Essas características garantem a continuidade da estrutura a termo implícita das taxas de juros futuras, condição essencial em contextos gerais de não-arbitragem financeira (Ang, Piazzesi e Wei 2006).

² (Chambers, Charlenton e Waldman 1984), (Diament 1993) e (Mansi e Phillips 2001) propuseram fórmulas funcionais para descrever a curva de rendimentos, porém, são pouco utilizadas na prática.

³ Broyden-Fletcher-Goldfarb-Shanno, método de otimização não-linear da classe Quasi-Newton.

para os parâmetros⁴. Atualmente, no mercado brasileiro, a Associação Brasileira das Entidades de Mercado Financeiro e de Capitais (AMBIMA) fornece as estimativas da ETTJ de acordo com a metodologia proposta por (Gimeno e Nave 2009).

Este artigo, considerando também métodos baseados em AEs, sugere um algoritmo de Evolução Diferencial (ED) para a estimação dos parâmetros das funções de NSS e obtenção da estrutura a termo da taxa de juros no mercado brasileiro. AEs são caracterizados por evoluir iterativamente uma população de soluções candidatas de um problema de otimização em busca das melhores soluções até que algum critério de parada seja satisfeito. A Evolução Diferencial pertence à classe dos algoritmos evolutivos que utilizam operadores bio-inspirados de crossover, mutação e seleção em uma população com o objetivo de otimizar uma função objetivo ao longo de sucessivas gerações.

Meta-heurística de busca introduzida por (Storn e Price 1997), os algoritmos de evolução diferencial apresentam uma elevada capacidade e notável performance em tratar problemas contínuos de otimização numérica e vêm sendo extensivamente explorados na literatura (Price, Storn e Lampinen 2006, Lampinen 2009). Em economia e finanças, algoritmos de ED também têm se mostrado como uma importante ferramenta na solução de problemas de otimização: na estimação de sofisticados modelos econométricos (Gilli, Maringer e Winker 2008, Gilli e Winker 2008), na seleção de modelos (Maringer e Meyer 2008, Maringer e Oyewumi 1997), e também na otimização de portfólios baseados em conjuntos não-convexos (Krink e Paterlini 2011, Krink, Mittnik e Paterlini 2009).

Dessa forma, com base em dados diários dos títulos públicos das Letras do Tesouro Nacional (LTNs), para o período de Janeiro de 2007 a Dezembro de 2009, foram obtidas as estimativas dos parâmetros das funções de NSS de acordo com um algoritmo de evolução diferencial proposto. Os resultados foram avaliados em comparação com a técnica clássica de mínimos quadrados não-lineares, em termos de métricas de erro quadrático e testes estatísticos.

Este artigo está organizado como segue. Após essa breve introdução, a Seção 2 apresenta os conceitos básicos da construção e estimação da ETTJ e o processo de otimização relacionado. A Seção 3 descreve as idéias principais do algoritmo de evolução diferencial, seguida pela Seção 4, composta pelos resultados e sua discussão. Por fim, a Seção 5 conclui o artigo e apresenta sugestões para trabalhos futuros.

2 Construção da Estrutura a Termo das Taxas de Juros

Esta seção apresenta as idéias centrais por trás da construção da curva de juros, assim como as funções de (Nelson e Siegel 1987) e (Svensson 1994), assim como os conceitos necessários para a formulação do problema de otimização para obtenção da curva zero-cupom.

2.1 Conceitos da Curva de Juros

A estrutura a termo da taxa de juros é definida como uma relação temporal entre as taxas nominais de juros de títulos zero-cupom e suas respectivas maturidades. O preço de um título público corresponde ao valor presente dos seus fluxos de caixa:

$$P = C \sum_{i=1}^m \delta(t_i) + R\delta(t_m) \quad (1)$$

em que P indica o preço do título, $\delta(t_i)$ a função de desconto, C o valor do cupom pago periodicamente em $t_i : i = 1, \dots, m$ e R o valor de face do título em t_m ⁵.

É conveniente considerar as taxas de juros continuamente capitalizadas, i.e., assumir que o desconto é realizado em períodos infinitesimais. Nesse caso, podemos definir a função de desconto como sendo:

$$\delta(t) = e^{-ty(t)} \quad (2)$$

⁴ (Fernández-Rodríguez 2006) também estimou a ETTJ por meio de algoritmos genéticos, porém, baseado em técnicas de *splines*.

⁵ Nos títulos zero-cupom, i.e., títulos que não pagam cupons, seu rendimento até a liquidação é igual a taxa *spot* para sua respectiva maturidade.

em que $y(t)$ é a taxa de juros *spot* com vencimento associado em t .

Conhecido os fatores de desconto $\delta(t_i)$, pode-se calcular as taxas futuras implícitas, ou seja, as taxas aplicadas a um determinado período específico:

$$r(t) = \frac{-\delta'(t)}{\delta(t)} \quad (3)$$

Obtém-se, então, a taxa média futura em um intervalo de maturidade definido $[m_1, m_2]$:

$$f(m_1, m_2) = \frac{1}{m_2 - m_1} \int_{m_1}^{m_2} r(t) dt \quad (4)$$

Dessa forma, a taxa *spot* com vencimento em m , é igual a taxa futura média no intervalo $[0, m]$:

$$y(m) = \frac{1}{m} \int_0^m \frac{-\delta'(t)}{\delta(t)} dt = \frac{-\ln(\delta(m))}{m} \quad (5)$$

Portanto, o preço de um título pode ser calculado da seguinte forma:

$$P = \sum_{i=1}^m C e^{-ty(i)} + R e^{-ty(i)} \quad (6)$$

O termo curva de juros, ou *yield curve*, é muito geral e corresponde, essencialmente, a curva que descreve a estrutura a termo das taxas de juros. Pode estar associado à curva das taxas de juros *spot* (curva zero-cupom), curva instantânea das taxas futuras, função de desconto, curva de amortização de rendimentos, e curva de rendimento nominal. Neste trabalho, será considerada a curva de juros zero-cupom, uma das construções de curva de juros mais utilizadas (Manousopoulos e Michalopoulos 2005).

2.2 Funções de Nelson-Siegel-Svensson

São diversos os métodos de estimação da curva de juros, organizados em duas categorias principais: métodos baseados em *splines* e métodos parcimoniosos. A primeira categoria aproxima a curva de juros via interpolação de funções polinomiais (cúbica, em geral), e por se tratar de um método mais geral, é capaz de capturar qualquer formato da curva de rendimentos. Os métodos parcimoniosos utilizam uma forma funcional única com poucos parâmetros para descrever a ETTJ, o que permite explorar o fato de que a curva de juros não apresenta uma forma arbitrária, mas que obedece restrições financeiras específicas, resultando numa representação mais simples e compacta⁶. Neste trabalho, para a estimação da curva de juros zero-cupom, consideramos dois modelos parcimoniosos amplamente utilizados, o modelo de (Nelson e Siegel 1987) e o modelo *Extended Nelson-Siegel* ou de (Svensson 1994), i.e., as funções Nelson-Siegel-Svensson.

O método parcimonioso desenvolvido por (Nelson e Siegel 1987), motivado pelo fato de que a curva de juros é monotônica, trata-se de um modelo paramétrico em que as taxas instantâneas futuras (r), em função da maturidade (m), apresenta a seguinte forma:

$$r(m) = \beta_1 + \beta_2 \cdot e^{\left(\frac{-m}{\tau_1}\right)} + \beta_3 \cdot \left(\frac{m}{\tau_1}\right) \cdot e^{\left(\frac{-m}{\tau_1}\right)} \quad (7)$$

em que β_1 , β_2 , β_3 e τ_1 são os parâmetros a serem estimados.

A Equação (7) é essencialmente uma função de Laguerre somada à uma constante, e possui características desejáveis para se capturar a forma da estrutura a termo. Uma questão fundamental é a existência dos limites da função:

$$\lim_{m \rightarrow \infty} r(m) = \beta_1 \quad (8)$$

$$\lim_{m \rightarrow 0} r(m) = \beta_1 + \beta_2 \quad (9)$$

⁶ Cada técnica possui suas vantagens e desvantagens, e embora não seja claro qual apresente melhor desempenho, modelos parcimoniosos têm sido mais utilizados por pesquisadores e participantes do mercado (Choudhry 2004).

o que permite a obtenção das taxas futuras instantâneas de longo e curto prazos, respectivamente.

A respectiva curva de rendimentos zero-cupom, que fornece a taxa à vista de um título zero-cupom com maturidade m , $z(m)$, é obtida integrando a função da taxa futura (Equação (7)):

$$z(m) = \frac{1}{m} \int_0^m r(x) dx = \beta_1 + \beta_2 \cdot \left(\frac{1 - e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} \right) + \beta_3 \cdot \left(\frac{1 - e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} - e^{-\frac{m}{\tau_1}} \right) \quad (10)$$

Os parâmetros possuem a seguinte interpretação: β_1 é a taxa de juros de longo prazo, $\beta_1 + \beta_2$ o valor da curva de curto prazo, enquanto β_3 e τ_1 determinam a forma em que ocorre a transição entre a taxa de curto e longo prazos, assim como a magnitude e direção do decaimento da curva⁷.

Uma vez que as taxas de juros *spot* são nominais e, por definição, estritamente positivas, para que a função faça sentido econômico, os parâmetros apresentam as seguintes restrições:

$$\beta_1 > 0, \quad \beta_1 + \beta_2 > 0 \quad e \quad \tau_1 > 0 \quad (11)$$

As duas primeiras inequações garantem que os externos da curva serão números positivos, enquanto a terceira inequação faz com que a função possua um limite finito para o termo de longo prazo.

Como uma extensão desse modelo, (Svensson 1994) adicionou um termo extra na função de Nelson-Siegel, para permitir uma segunda curvatura (*hump*) na estrutura a termo. Essa nova versão, denominada de modelo *Extended Nelson-Siegel* ou modelo de Svensson é capaz de capturar uma maior variedade de formatos da curva de juros, especialmente em ambientes de crises monetárias. A curva instantânea das taxas futuras é definida como:

$$r(m) = \beta_1 + \beta_2 \cdot e^{\left(\frac{-m}{\tau_1}\right)} + \beta_3 \cdot \left(\frac{m}{\tau_1}\right) \cdot e^{\left(\frac{-m}{\tau_1}\right)} + \beta_4 \cdot \left(\frac{m}{\tau_2}\right) \cdot e^{\left(\frac{-m}{\tau_2}\right)} \quad (12)$$

De forma similar, ao integrar a função da taxa futura, a respectiva curva de juros zero-cupom é:

$$z(m) = \beta_1 + \beta_2 \cdot \left(\frac{1 - e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} \right) + \beta_3 \cdot \left(\frac{1 - e^{-\frac{m}{\tau_1}}}{\frac{m}{\tau_1}} - e^{-\frac{m}{\tau_1}} \right) + \beta_4 \cdot \left(\frac{1 - e^{-\frac{m}{\tau_2}}}{\frac{m}{\tau_2}} - e^{-\frac{m}{\tau_2}} \right) \quad (13)$$

Um valor positivo (negativo) para o parâmetro β_4 implica em um “U” (“U” invertido) adicional na curva, cuja posição é determinada por τ_2 . Se $\tau_2 > \tau_1$ ($\tau_2 < \tau_1$), o efeito do último termo desaparece mais rápido (devagar), o que implica que será observado antes (depois) a forma determinada por β_3 e τ_1 ⁸. Além disso, β_4 e τ_2 determinam a posição e magnitude, respectivamente, de uma segunda curvatura na ETTJ. Para se verificar sentido econômico, as restrições aos parâmetros são as mesmas que no modelo Nelson-Siegel, incluindo $\tau_2 > 0$, o que garante um limite finito de longo prazo.

2.3 Processo de Otimização

Para estimar a curva de juros zero-cupom são considerados títulos públicos transacionados no mercado e então se determina os parâmetros β_i e τ_j que minimizam uma medida de erro entre as taxas ou preços observados no mercado e suas respectivas estimativas, obtidas pelo modelo considerado. A escolha entre preços ou taxas dos títulos para determinação da função objetivo (ou custo) do processo de otimização não é definitiva. Para selecionar o método apropriado, deve-se levar em consideração o objetivo da estimação da curva de juros. Por exemplo, se a curva estimada for utilizada como indicador de política monetária, i.e., o foco são as taxas de juros, então o processo de minimização deve considerar as taxas. Por outro lado, quando a curva obtida é destinada para precificação de títulos, então a otimização em função dos preços deve ser considerada (Svensson 1994). Neste trabalho, serão utilizados os preços dos títulos na determinação da função objetivo.

O critério de otimização determina o conjunto de parâmetros $\Omega = \{\beta_i, \tau_j\}$ que minimiza a diferença quadrática entre os preços dos títulos observados no mercado (P_k) e os preços obtidos pelo modelo (\hat{P}_k), ou seja, a função custo $F(\Omega)$ é definida da seguinte forma:

⁷ (Diebold e Li 2006) afirmam que os parâmetros β_1 , β_2 e β_3 podem ser interpretados como fatores de nível, inclinação e curvatura, respectivamente.

⁸ Quando $\beta_4 = 0$ ou $\tau_1 = \tau_2$ a curva de Svensson se torna a função de Nelson-Siegel.

$$\text{Min } F(\Omega) = \text{Min} \sum_{k=1}^N (P_k - \hat{P}_k)^2 \cdot \frac{1}{\sqrt{D_k}} \quad (14)$$

em que N é o número de títulos transacionados diariamente e D_k a *duration* do título k ⁹.

A ponderação do erro quadrático dos preços pela *duration* objetiva controlar a heterocedasticidade, uma vez que uma pequena mudança nos preços dos títulos de maturidade curta implica em um grande movimento na taxa curta, enquanto que para títulos de prazos mais longos ocorre o oposto. Portanto, a minimização penaliza mais os erros de títulos de curto prazos ao se ponderar pelo inverso da *duration* (Gimeno e Nave 2009).

No processo de otimização é necessário determinar os valores iniciais dos parâmetros β_i e τ_j . A solução de um problema de programação não-linear é muito sensível à determinação do ponto inicial, uma vez que é crucial para a acurácia da curva resultante. Os parâmetros das funções Nelson-Siegel-Svensson possuem uma interpretação financeira e isso permite a escolha dos valores iniciais mais eficientes para a convergência do algoritmo. Neste trabalho, foi adotado o seguinte critério para determinação desses parâmetros, de acordo com (Manousopoulos e Michalopoulos 2009):

$$\beta_1 = \frac{1}{M} \sum_{i=1}^M y_i, \beta_2 = y_s - \beta_1, \beta_3 = \beta_4 = 0, \tau_1 = \tau_2 = 1 \quad (15)$$

na qual M se refere aos títulos de maturidades mais longas, usualmente 2 ou 3 anos, e s representa a taxa referente ao título de menor maturidade.

Dessa forma, aqui a curva de juros zero-cupom é estimada de acordo com as funções de Nelson-Siegel e Svensson por meio da técnica de mínimos quadrados não-lineares e de acordo com um modelo de algoritmo de evolução diferencial, conforme será descrito¹⁰.

3 Algoritmo de Evolução Diferencial

3.1 Computação Evolutiva

A computação evolutiva se inspira em princípios da teoria da evolução e seleção natural e utiliza modelos destes processos naturais para a solução de problemas de otimização, controle, classificação. Seus principais ramos são: Algoritmos Genéticos, Estratégias Evolutivas, Programação Evolutiva, Programação Genética e Evolução Diferencial.

As técnicas baseadas em Algoritmos Genéticos (AGs), pioneiras na área de computação evolutiva, foram introduzidas por (Holland 1975) e se resumem numa técnica de busca, inspirada na teoria da evolução. Desde então, essa metodologia se mostrou eficiente na resolução de inúmeros problemas de otimização, inclusive em economia e finanças (Mckee e Lensberg 2002, Roudier 2007, Lin, Cao, Wang e Zhang 2004). O mecanismo de funcionamento de um algoritmo genético parte, inicialmente, de uma população inicial de indivíduos, ou cromossomos, gerados de forma aleatória, assim como se é verificado na natureza. Em um processo de otimização, cada cromossomo representa uma possível solução ao problema em questão. Essa população evolui por meio da criação de novas gerações, sendo que a cada geração são indivíduos dessa população selecionados e, por meio de operadores de reprodução e mutação, estes determinam os novos indivíduos da geração seguinte.

A cada geração são selecionados os indivíduos mais aptos. A aptidão está associada à solução do problema de otimização; uma função objetivo (ou função de *fitness*) seleciona os indivíduos com o melhor valor. Portanto, os indivíduos que se são as melhores soluções ao problema (mais aptos) são selecionados

⁹ A não-linearidade do modelo de otimização é imposta pelos parâmetros τ 's. Muitos trabalhos, como (Diebold e Li 2006), fixam arbitrariamente esses parâmetros e obtêm as estimativas de acordo com o método de mínimos quadrados ordinários, porém, esse procedimento implica numa limitação, uma vez que não se verifica um procedimento para se determinar um valor adequado à esses parâmetros. Neste trabalho consideramos todos os parâmetros variáveis no processo de otimização.

¹⁰ Para detalhes da metodologia de mínimos quadrados não-lineares ver (Strutz 2010).

para reproduzir, recombinar e sofrer mutação para então resultar na nova geração. Ao longo das gerações (iterações), a população em geral converge para a solução ótima do problema¹¹.

Nos métodos tradicionais de otimização, ao partir de um único ponto inicial, verifica-se o risco da falsa convergência, i.e., solução ser um mínimo local, principalmente quando se considera funções objetivo não-lineares. Entretanto, os modelos baseados em computação evolutiva asseguram que nenhum ponto do espaço tem probabilidade zero de ser examinado, o que aumenta as chances de se obter como solução um ótimo global.

3.2 Evolução Diferencial

O algoritmo de Evolução Diferencial (ED) proposto por (Storn e Price 1997) nasceu da idéia de se utilizar um vetor de diferenças para perturbar o vetor da população. Trata-se de uma abordagem aprimorada dos AGs para processos de otimização e é essencialmente um otimizador estocástico de funções baseado numa população de possíveis soluções. A idéia do algoritmo de ED consiste em um esquema de geração de vetores de parâmetros experimentais que irão perturbar o vetor da população.

(Mitchell 1998) afirma que os modelos baseados em computação evolutiva, em geral, possuem quatro componentes principais: inicialização, seleção, crossover e mutação. Esses componentes podem ser aplicados de diferentes maneiras de acordo com a técnica evolutiva utilizada e o problema considerado. A seguir eles serão descritos para o caso do algoritmo de evolução diferencial, considerando a aplicação proposta neste trabalho.

3.2.1 Inicialização

O método de evolução diferencial é um modelo de busca paralelo que utiliza um total de NP vetores de parâmetros como população em cada geração G , isto é:

$$x_{i,G} \quad i = 1, 2, \dots, NP \quad (16)$$

Durante o processo de otimização NP é mantido constante. Na ausência de conhecimento do espaço de busca, a população inicial é gerada aleatoriamente, considerando uma distribuição de probabilidade uniforme. Portanto, com uma solução semente (nominal), a população inicial é gerada somando desvios aleatoriamente distribuídos à solução semente.

No caso da estimação da curva de juros, de acordo com as funções de Nelson-Siegel-Svensson, é necessário obter os parâmetros que as definem, portanto, cada indivíduo da população no algoritmo de evolução diferencial é representado por uma possível solução do problema em questão. Representamos o indivíduo i da população de tal forma que:

$$x_i = \Omega_i = (\beta_{1j}, \beta_{2j}, \beta_{3j}, \beta_{4j}, \tau_{1j}, \tau_{2j}) \quad (17)$$

Então, cada indivíduo i da população é um conjunto dos parâmetros que descrevem as funções de Nelson-Siegel e Svensson. A população inicial, de tamanho NP é gerada aleatoriamente da seguinte forma:

$$\beta_{ij} = \beta_i^* + \epsilon_{ij}, \quad i = 1, \dots, 4, \quad j = 1, \dots, NP - 1 \quad (18)$$

$$\tau_{hj} = \tau_h^* + \zeta_{hj}, \quad h = 1, 2 \quad j = 1, \dots, NP - 1 \quad (19)$$

onde β_i^* e τ_h^* são os valores com os quais a população inicial é criada (solução semente), $\epsilon_{ij} \approx N(0, \sigma_{\beta_i})$ e $\zeta_{hj} \approx N(0, \sigma_{\tau_h})$ são valores aleatórios¹².

Os valores sementes dos parâmetros para geração da população inicial foram definidos conforme a Expressão (15), que leva em consideração as características dos títulos utilizados na estimação e conceitos do mercado financeiro. Deve-se notar que os valores iniciais não interferem tanto no resultado do

¹¹ Os demais ramos da computação evolutiva seguem, em geral, esses princípios básicos propostos nos algoritmos genéticos, porém cada qual com sua especificidade.

¹² Para fins de simplificação, $\sigma_{\beta_i} = \sigma_{\tau_h} = \sigma$, que descreve a variabilidade da população inicial.

algoritmo de evolução diferencial, ao contrário dos algoritmos de otimização tradicionais, mas nesse caso reduzem o tempo de busca. As restrições quanto aos parâmetros são as mesmas já mencionadas na Seção 2, para que eles possuam sentido econômico, assim, considerou-se apenas os indivíduos factíveis.

3.2.2 Mutação

A idéia principal por trás do algoritmo de evolução diferencial é gerar um novo vetor de parâmetros pela adição da diferença ponderada entre dois membros da população à um terceiro membro. Esse mecanismo é considerado como um operador de mutação no caso do algoritmo de evolução diferencial.

Para cada vetor $x_{i,G}$, $i = 1, 2, \dots, NP$, na geração G , denominado *target vector*, um novo vetor $v_{i,G+1}$, é gerado de acordo com a seguinte expressão:

$$v_{i,G+1} = x_{r_1,G} + F(x_{r_2,G} - x_{r_3,G}) \tag{20}$$

onde $r_1, r_2, r_3 \in [0, NP]$ são índices mutuamente distintos e também diferentes do índice i e F uma constante real.

Os inteiros r_1, r_2, r_3 são escolhidos aleatoriamente no intervalo $[0, NP]$, e $F \in [0, 2]$ controla o tamanho do passo a ser dado na direção definida pelo vetor diferença $(x_{r_2,G} - x_{r_3,G})$. A Figura 1 apresenta um exemplo do processo de mutação para a geração de um novo indivíduo, considerando uma função objetivo bidimensional.

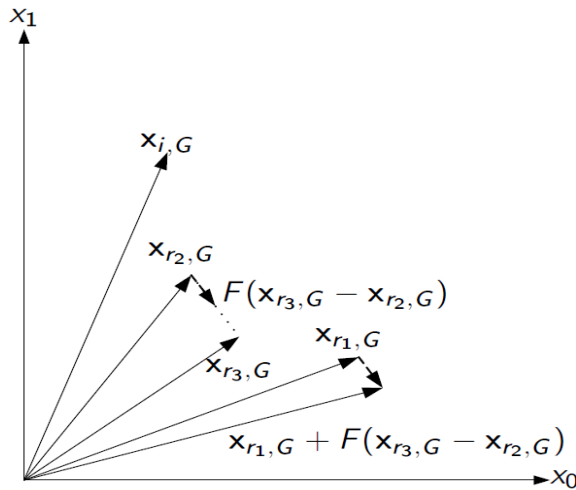


Fig. 1: Processo de Mutação no Algoritmo de Evolução Diferencial.

A mutação permite indivíduos em todo espaço de busca, mas enfatiza aqueles que fornecem um melhor valor da função objetivo.

Inúmeras variações do algoritmo de evolução diferencial clássico foram propostos na literatura, e na maioria das vezes, são distintos quanto ao operador de mutação. Além de operador apresentado, serão avaliados outras duas propostas de operadores de mutação. A primeira delas, *ED/best* possui a seguinte forma:

$$v_{i,G+1} = x_{best,G} + F(x_{r_2,G} - x_{r_1,G}) \tag{21}$$

em que $x_{best,G}$ representa o indivíduo que fornece o melhor resultado para a função objetivo em uma determinada geração G .

O outro operador, denominado *ED/target-to-best*, é o seguinte:

$$v_{i,G+1} = x_{i,G} + F(x_{best,G} - x_{i,G}) + F(x_{r_2,G} - x_{r_3,G}) \tag{22}$$

O operador descrito inicialmente, na Expressão (20), é denominado *ED/rand*, pois a mutação ocorre entre indivíduos selecionados randomicamente.

3.2.3 Crossover

Para aumentar a diversidade dos vetores de parâmetros, o algoritmo de evolução diferencial utiliza um processo similar ao crossover. Considere um determinado *target vector* $x_{i,G}$ e seu respectivo vetor mutado $v_{i,G+1}$, obtido da Expressão (20), o vetor $u_{i,G+1} = (u_{1i,G} \ u_{2i,G} \ \dots \ u_{NPi,G})$, denominado *trial vector*, é obtido da seguinte forma:

$$u_{ji,G+1} = \begin{cases} v_{ji,G+1}, & \text{se } r_j \leq CR \text{ ou } j = l_i \\ x_{ji,G}, & \text{se } r_j > CR \text{ e } j \neq l_i \end{cases} \quad (23)$$

onde r_j é uma variável aleatória com distribuição uniforme no intervalo $[0, 1]$, tal que $j \in [1, NP]$, $CR \in [0, 1]$, definido como taxa de crossover, é uma constante definida pelo usuário, e $l_i \in [0, NP]$ um índice escolhido aleatoriamente.

A função de l_i é garantir que o *trial vector* $u_{i,G+1}$ receba pelo menos uma componente de $v_{i,G+1}$. Cada vetor da população atual deve ser usado uma vez como *trial vector*. A Figura 2 exemplifica o processo de geração do *trial vector*, representando a operação de crossover em algoritmos de evolução diferencial.

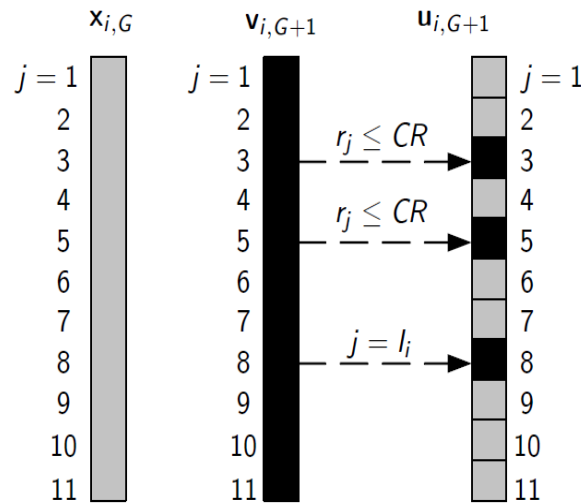


Fig. 2: Processo de Crossover no Algoritmo de Evolução Diferencial.

3.2.4 Seleção

Depois das etapas de mutação e crossover, de acordo com o *target vector* $x_{i,G}$ e seu respectivo *trial vector* $u_{i,G+1}$, no qual todos os NP vetores serviram como *target vector*, a seleção dos vetores que serão preservados para a próxima geração é feita considerando o seguinte critério:

- Se $F(u_{i,G+1}) < F(x_{i,G})$, o novo vetor de parâmetros da geração $G + 1$ é $x_{i,G+1} = u_{i,G+1}$ ¹³;
- Caso contrário $x_{i,G+1} = x_{i,G}$.

Em que $F(\cdot)$ representa o valor da função custo considerada no algoritmo de otimização. Assim, esse processo de seleção avalia os indivíduos gerado ao longo de uma geração atual e com os indivíduos da geração anterior, sendo que a nova geração será composta pelos melhores indivíduos, i.e., aqueles que fornecem melhores resultados para o problema em questão.

Os processos de mutação, crossover e seleção se repetem de acordo com o critério de parada escolhido, que em geral, corresponde ao número máximo de gerações.

¹³ Considerando um processo de minimização.

3.2.5 Algoritmo

Podemos sintetizar o algoritmo de evolução diferencial através dos seguintes passos:

Algoritmo de Evolução Diferencial	
1	Iniciar Ω_1 como semente inicial
2	Gerar aleatoriamente a população inicial Ω de soluções ao redor de Ω_1
3	Iniciar as gerações, Iteração = 1
4	while Iteração < Gerações do
5	for $i = 1$ até NP do
6	Mutar os indivíduos: $v_{i,G+1} \leftarrow \text{mutação}(x_{i,G}, F)$
7	Realizar o Crossover: $u_{i,G+1} \leftarrow \text{crossover}(x_{i,G}, v_{i,G+1}, CR)$
8	end for
9	Computar o <i>fitness</i> dos indivíduos: $fit_u \leftarrow f(u)$
10	for $i = 1$ até NP do
11	Selecionar os indivíduos:
12	if $fit_u(i) < fit_x(i)$ then
13	$x_{i,G+1} \leftarrow u_{i,G+1}$
14	else
15	$x_{i,G+1} = x_{i,G}$
15	end if
16	end for
17	Iteração = Iteração + 1
18	end while

O algoritmo de evolução diferencial proposto neste trabalho utiliza cinco parâmetros que devem ser escolhidos pelo usuário: NP , CR , F , σ e G . Apenas o número de gerações G , e o desvio-padrão σ , para inicialização da população, serão fixados em $G = 400$ e $\sigma = 0,5$. Os demais parâmetros serão avaliados para determinar a sensibilidade da função objetivo aos seus valores, e então obter aqueles que apresentam melhores resultados. A função objetivo ou de *fitness* considerada corresponde à Equação (14), descrita na seção anterior.

O método de evolução diferencial se torna atrativo pela capacidade de lidar com funções objetivo não-lineares, não-diferenciáveis e multimodais. Além disso, é passível de paralelização, é acessível, permite o auto-ajuste do passo de adaptação, assim como possui boas propriedades de convergência (Storn e Price 1997). De acordo com suas características e peculiaridades, essa metodologia mostra-se como uma boa alternativa para superar as limitações dos métodos tradicionais de estimação da curva de juros, considerando modelos parcimoniosos.

4 Resultados e Discussão

A ETTJ nominal foi estimada diariamente para o período de Janeiro de 2007 a Dezembro de 2009. Os dados utilizados foram os preços indicativos dos títulos públicos federais, divulgados pela Associação Brasileira das Entidades de Mercado Financeiro e de Capitais. Os títulos prefixados considerados foram as Letras do Tesouro Nacional (LTNs), que são títulos zero-cupom, emitidos pelo Tesouro Nacional para todos os vencimentos disponíveis.

Inicialmente, comparou-se o desempenho da técnica de evolução diferencial considerando operadores de mutação distintos, ou seja, as estratégias: *ED/ rand*, *ED/ best* e *ED/ target-to-best*, conforme definidas na Seção 3. A Tabela 1 apresenta a variabilidade dos parâmetros estimados, medida pelo desvio-padrão, e o erro quadrático médio (EQM)¹⁴ entre os preços de mercado e os preços obtidos pelas técnicas consideradas, de acordo com a seguinte parametrização: $NP = 20$, $CR = 0,9$ e $F = 1,25$.

¹⁴ Define-se $EQM = \frac{1}{N} \sum_{k=1}^N (P_k - \hat{P}_k)^2$.

Tab. 1: Variabilidade dos parâmetros das funções de NSS e EQM dos preços de mercado e estimados de acordo com as técnicas de evolução diferencial empregadas.

Função de Nelson-Siegel							
Estratégias	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	$EQMP$
<i>ED/rand</i>	0,0355	0,1000	0,2380	—	12,3142	—	0,4857
<i>ED/best</i>	0,0419	0,1057	0,2436	—	12,5910	—	0,8012
<i>ED/target-to-best</i>	0,0391	0,0586	0,2323	—	12,5714	—	0,4659
Função de Svensson							
Estratégias	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	$EQMP$
<i>ED/rand</i>	0,0320	0,1532	0,2383	0,2658	11,3114	11,7670	0,4342
<i>ED/best</i>	0,0306	0,1564	0,2418	0,2716	11,7058	11,8443	0,4402
<i>ED/target-to-best</i>	0,0326	0,1029	0,2108	0,2322	11,5573	11,1892	0,3632

De acordo com a Tabela 1, tanto para a função de Nelson-Siegel quanto para a de Svensson, o algoritmo de evolução diferencial que utiliza como operador de mutação a direção do melhor indivíduo da população para gerar novos indivíduos, *ED/target-to-best*, gerou o menor valor para a função objetivo, garantindo maior aderência do modelo estimado aos dados de mercado. Além disso, essa mesma estratégia, em geral, forneceu estimativas menos voláteis para os fatores das funções de NSS, o que evidencia a melhor adequação da curva aos princípios financeiros por trás da estimação da ETTJ¹⁵.

Dessa forma, foi considerado o modelo de algoritmo de evolução diferencial *ED/target-to-best* para estimar a ETTJ de acordo com as funções de NSS. Variou-se, então, os parâmetros que definem a técnica de evolução diferencial para se obter a parametrização que forneça os melhores resultados na construção da curva de juros. Mantendo constante $CR = 0,9$ e $F = 1,25$, a Tabela 2 apresenta os resultados obtidos de acordo com a variação no parâmetro do tamanho da população.

Tab. 2: Variabilidade dos parâmetros das funções de NSS e EQM dos preços de mercado e estimados de acordo com a variação do tamanho da população no algoritmo *ED/target-to-best*.

Função de Nelson-Siegel							
NP	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	$EQMP$
20	0,0391	0,0586	0,2323	—	12,5714	—	0,4659
40	0,0384	0,0574	0,2298	—	12,6477	—	0,4536
60	0,0402	0,0563	0,2351	—	12,5698	—	0,3814
80	0,0396	0,0625	0,2425	—	12,7444	—	0,4499
100	0,0412	0,0688	0,2386	—	12,7920	—	0,4759
Função de Svensson							
NP	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	$EQMP$
20	0,0326	0,1029	0,2108	0,2322	11,5573	11,1892	0,3632
40	0,0329	0,1014	0,2063	0,2344	11,3687	10,9865	0,4019
60	0,0358	0,1125	0,1977	0,2286	12,7894	10,7548	0,5524
80	0,0598	0,1099	0,2159	0,2217	11,6531	11,1197	0,5756
100	0,0571	0,1035	0,2208	0,2303	11,4855	11,2684	0,6611

Para a função de Nelson-Siegel, um relativo aumento no tamanho da população de indivíduos resultou num melhoramento da função objetivo, mas quando esse número cresce muito, a aderência da curva de juros aos dados de mercado diminui, de forma que, com 60 indivíduos na população, obtém-se as melhores estimativas dos fatores da ETTJ (Tabela 2). Por outro lado, ao se considerar a função de Svensson, o

¹⁵ Na estimação da curva de juros, de acordo com as funções de Nelson-Siegel-Svensson, é necessário que se obtenha estimativas dos parâmetros que forneçam uma curva mais aderente aos preços dos títulos considerados e que ao mesmo tempo tenham baixa variabilidade, uma vez que dada a interpretação econômica dos parâmetros, eles revelam as expectativas do mercado quanto ao futuro das taxas de curto e longo prazos e, desconsiderando momentos de instabilidade no mercado monetário, essas expectativas não devem apresentar alta variabilidade (Gimeno e Nave 2009).

parâmetro NP é inversamente proporcional ao desempenho do algoritmo, tal que $NP = 20$ produz os melhores resultados. Contudo, deve-se salientar que o custo computacional cresce de acordo com o aumento do número de indivíduos na população, e como a estimação da ETTJ é um processo que requer dinamismo, um número relativamente baixo para NP torna o algoritmo mais parcimonioso e adequado para o objetivo aqui proposto.

Mantendo $F = 1,25$ e associando os valores $NP = 60$ para a função de Nelson-Siegel e $NP = 20$ para a função de Svensson, a Tabela 3 indica a sensibilidade dos resultados ao se variar o parâmetro que define a taxa de crossover (CR). Pode-se perceber que a performance do algoritmo de evolução diferencial diminui quando se considera uma taxa baixa de crossover para estimação das funções de NSS, fato que se explica pelo motivo de que a operação de crossover permite fornecer maior variabilidade aos indivíduos que irão compor a próxima geração, ampliando o espaço de busca do algoritmo, de tal forma que facilita a obtenção do ponto de ótimo global.

Tab. 3: Variabilidade dos parâmetros das funções de NSS e EQM dos preços de mercado e estimados de acordo com a variação da taxa de crossover no algoritmo $ED/target-to-best$.

Função de Nelson-Siegel							
CR	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	EQM_P
0.3	0,0440	0,0666	0,2396	—	11,7841	—	0,6656
0.5	0,0398	0,0854	0,2388	—	11,9961	—	0,5654
0.7	0,0423	0,0642	0,2449	—	11,9475	—	0,5377
0.9	0,0402	0,0563	0,2351	—	12,5698	—	0,3814
Função de Svensson							
CR	$\hat{\sigma}_{\beta_1}$	$\hat{\sigma}_{\beta_2}$	$\hat{\sigma}_{\beta_3}$	$\hat{\sigma}_{\beta_4}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau_2}$	EQM_P
0.1	0,0425	0,1125	0,2214	0,2487	11,8741	12,6581	0,6885
0.3	0,0356	0,1247	0,2377	0,2214	11,6602	11,1986	0,5944
0.5	0,0339	0,1099	0,2184	0,2398	12,5774	11,1963	0,4799
0.7	0,0311	0,1157	0,2145	0,2364	12,0087	10,9855	0,2168
0.9	0,0326	0,1029	0,2108	0,2322	11,5573	11,1892	0,3632

Finalmente, a Figura 4 mostra a variação do erro quadrático médio (EQM) dos preços dos títulos em função da variação do parâmetro que controla o passo a ser dado na direção do vetor diferença entre os indivíduos da população (F), considerando os demais parâmetros obtidos anteriormente que forneceram as melhores estimativas dos fatores das funções de NSS.

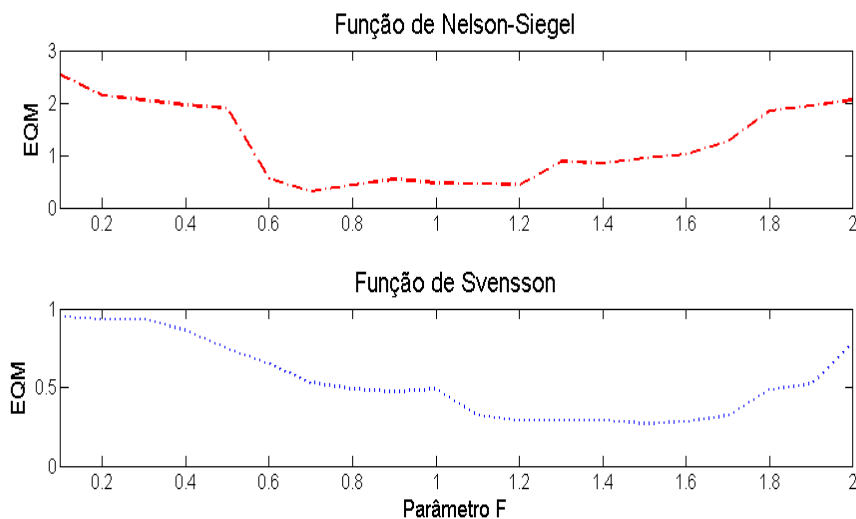


Fig. 3: Variação do EQM para as Funções de NSS de acordo com a variação do parâmetro F .

Pode-se perceber relativa sensibilidade dos resultados de acordo com a variação do parâmetro F , de forma que os resultados pioram com a aproximação dos limites de F .

Foram obtidas, então, as estimativas dos parâmetros das funções de NSS de acordo com a técnica de mínimos quadrados não-lineares (MQNL) e com o algoritmo de evolução diferencial (ED) proposto. A Tabela 4 apresenta as estatísticas descritivas dos parâmetros das funções de Nelson-Siegel-Svensson de acordo com as técnicas de otimização empregadas. Apesar da similaridade dos resultados, quando obtidos por meio do algoritmo de evolução diferencial, os parâmetros das funções de NSS apresentam menor variabilidade, o que garante maior robustez à técnica proposta¹⁶.

Tab. 4: Estatísticas Descritivas dos parâmetros das funções de NSS estimados de acordo com as técnicas de MQNL e ED.

Função de Nelson-Siegel						
Evolução Diferencial						
Parametros	Média	Desv. Pad	Mínimo	Máximo	$\hat{\rho}^1$	ADF ¹
$\hat{\beta}_{1,t}$	0,0402	0,0391	1,19E-06	0,1232	0,7241	-5,5290
$\hat{\beta}_{2,t}$	0,0563	0,0586	5,92E-08	0,9801	0,2360	-16,5152
$\hat{\beta}_{3,t}$	0,2351	0,2323	2,45E-06	0,9996	0,9354	-3,4175
$\hat{\beta}_{4,t}$	-	-	-	-	-	-
$\hat{\tau}_{1,t}$	12,5698	12,5714	4,18E-09	30	0,9218	-4,1502
$\hat{\tau}_{2,t}$	-	-	-	-	-	-
MQNL						
Parametros	Média	Desv. Pad	Mínimo	Máximo	$\hat{\rho}^1$	ADF ¹
$\hat{\beta}_{1,t}$	0,0688	0,0432	7,58E-10	0,1265	0,9095	-3,2060
$\hat{\beta}_{2,t}$	0,0592	0,0425	1,35E-13	0,1242	0,9180	-4,4969
$\hat{\beta}_{3,t}$	0,2100	0,2200	5,84E-07	0,9961	0,9308	-3,9414
$\hat{\beta}_{4,t}$	-	-	-	-	-	-
$\hat{\tau}_{1,t}$	13,3026	11,4389	0,0279	30	0,9350	-3,5028
$\hat{\tau}_{2,t}$	-	-	-	-	-	-
Função de Svensson						
Evolução Diferencial						
Parametros	Média	Desv. Pad	Mínimo	Máximo	$\hat{\rho}^1$	ADF ¹
$\hat{\beta}_{1,t}$	0,0311	0,0326	1,48E-06	0,1309	0,0984	-6,4875
$\hat{\beta}_{2,t}$	0,1157	0,1029	3,98E-05	0,9933	0,2158	-7,2541
$\hat{\beta}_{3,t}$	0,2145	0,2180	1,47E-05	0,9947	0,9584	-5,1974
$\hat{\beta}_{4,t}$	0,2364	0,2322	2,06E-06	0,9996	0,9632	-8,0955
$\hat{\tau}_{1,t}$	12,0087	13,2586	5,69E-06	29,9999	0,8962	-4,8461
$\hat{\tau}_{2,t}$	10,9855	16,3296	0,0002	16,3296	0,9211	-6,3371
MQNL						
Parametros	Média	Desv. Pad	Mínimo	Máximo	$\hat{\rho}^1$	ADF ¹
$\hat{\beta}_{1,t}$	0,0602	0,0658	1,19E-05	0,1269	0,1223	-6,8474
$\hat{\beta}_{2,t}$	0,0701	0,1326	4,01E-06	0,8659	0,2478	-5,2014
$\hat{\beta}_{3,t}$	0,1456	0,2411	1,69E-06	0,9922	0,9955	-6,6179
$\hat{\beta}_{4,t}$	0,1798	0,2698	3,15E-07	0,9840	0,9547	-7,1020
$\hat{\tau}_{1,t}$	13,2587	14,2588	5,44E-05	30	0,8882	-6,0009
$\hat{\tau}_{2,t}$	14,1458	15,8475	0,0008	17,9587	0,9025	-7,3147

¹ $\hat{\rho}$ representa o indicador de auto-correlação e ADF o teste de raiz unitária Dickey-Fuller.

¹⁶ Além disso, a estatística ADF, para ambas as metodologias de otimização, indica a rejeição da hipótese nula de raiz unitária (Tabela 4).

De acordo com a análise de sensibilidade dos parâmetros do algoritmo de evolução diferencial considerado, a Tabela 5 resume os resultados de parametrização e o EMQ obtidos para os fatores da curva de juros de acordo com as funções de NSS.

Tab. 5: Parametrização do algoritmo $ED/target-to-best$ para estimação dos parâmetros das funções de NSS.

Parâmetros	Função de Nelson-Siegel	Função de Svensson
NP	60	20
CR	0,90	0,70
F	0,75	1,50
σ	0,50	0,50
G	400	400
$EQMP$	0,3814	0,2168

A estabilidade dos parâmetros pode ser avaliada na Figura 4, que mostra os componentes de longo e curto prazos (β_1 e $\beta_1 + \beta_2$) das funções de NSS obtidos por meio da metodologia de MQNL e de evolução diferencial. Confirmam-se os resultados da Tabela 4, em que a técnica de evolução diferencial é capaz de gerar uma menor variabilidade aos componentes das funções de Nelson-Siegel e Svensson.

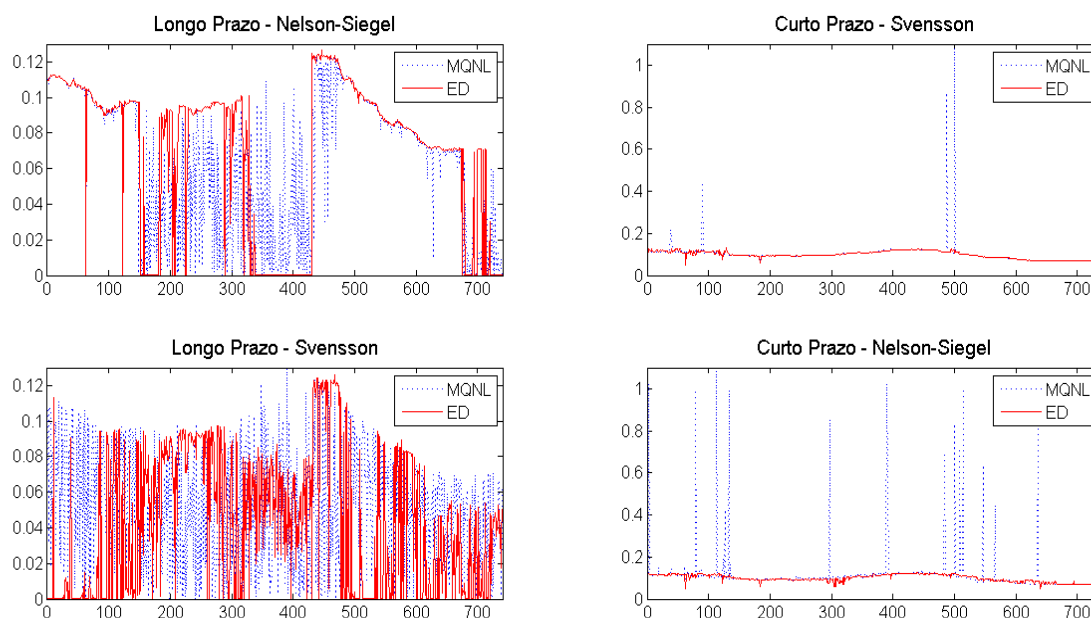


Fig. 4: Evolução dos componentes de longo (β_1) e curto ($\beta_1 + \beta_2$) prazos das funções de NSS estimadas de acordo com o método de MQNL e ED.

O erro quadrático médio entre os preços de mercado das LTNs e os preços estimados desses títulos obtidos por meio das abordagens avaliadas são mostrados na Tabela 6.

Tab. 6: EQM entre os preços dos títulos de mercado e estimados por meio das funções de NSS via MQNL e o algoritmo $ED/target-to-best$.

	Nelson-Siegel - MQNL	Nelson-Siegel - ED	Svensson - MQNL	Svensson - ED
$EQMP$	0,8012	0,3814	0,6402	0,2168

A técnica de evolução diferencial apresentou menores valores para o EQM na estimação das funções de Nelson-Siegel e Svensson, se comparado com a técnica de MQNL. Para confirmar tais resultados,

a Figura 5 apresenta a evolução da diferença entre os preços de mercado e os preços estimados de acordo com as técnicas consideradas para as funções de NSS. Pode-se perceber que os erros obtidos no algoritmo de evolução diferencial concentram-se mais em torno de zero, revelando maior precisão dessa abordagem na estimação das funções de NSS.

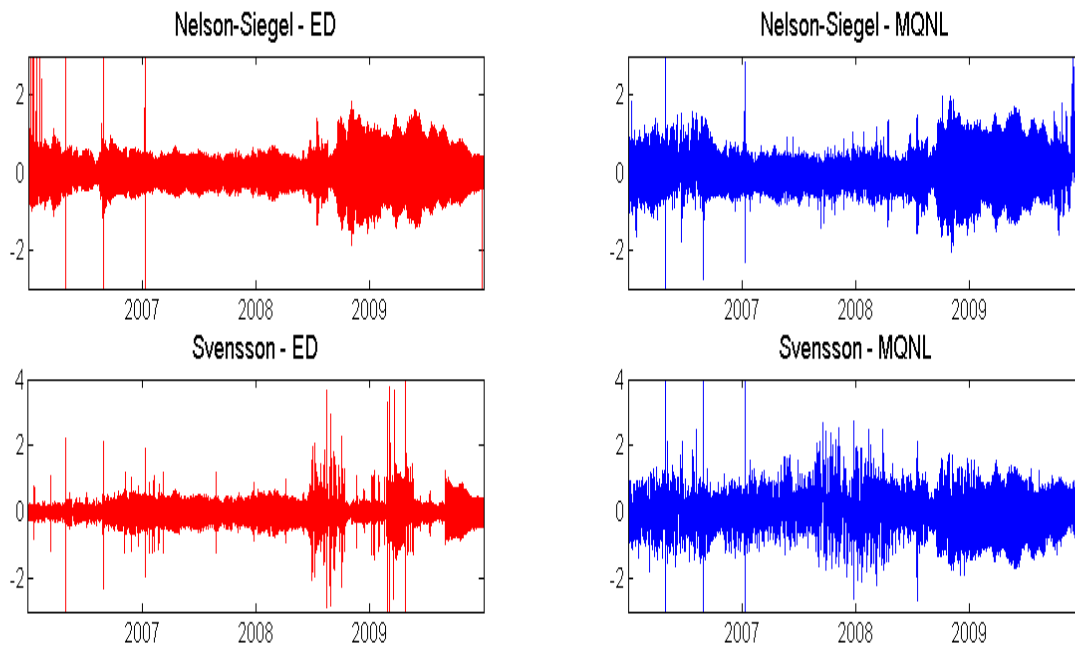


Fig. 5: Diferença entre os preços estimados e os preços de mercado para as funções de NSS de acordo com as técnicas de MQNL e ED.

Além da precisão, para avaliar a robustez dos resultados, levamos em consideração a comparação dos modelos em termos de significância estatística. Foram aplicados, então, um teste paramétrico e outro não-paramétrico para verificar se existe diferença estatisticamente significativa entre os resultados dos modelos considerados.

O teste paramétrico, denominado de teste de Morgan-Granger-Newbold (teste *MGN*), baseado inicialmente no trabalho de (Granger e Newbold 1977), é recomendado para avaliar modelos competitivos e é amplamente aplicado quando se relaxa a assunção de não existência de correlação entre os erros dos modelos. O valor da estatística para esse teste pode ser calculado como:

$$MGN = \frac{\hat{\rho}_{s,d}}{(1 - \hat{\rho}_{s,d})^{\frac{1}{2}}} (N - 1)^{\frac{1}{2}} \quad (24)$$

em que $\hat{\rho}_{sd}$ é o coeficiente de correlação estimado entre $s = e_{MQNL} + e_{ED}$ e $d = e_{MQNL} - e_{ED}$, tal que e_{MQNL} e e_{ED} correspondem aos erros dos modelos de MQNL e ED respectivamente¹⁷, e N representa o número de títulos na amostra. A estatística para o teste MGN tem distribuição *t-Student* com $(N - 1)$ graus de liberdade. Para este teste, se as previsões são igualmente acuradas, entre os modelos competitivos, então a correlação entre s e d será zero (hipótese nula).

Por fim, o teste não-paramétrico *SIGN* (*Significance Test*), derivado do trabalho de (Lehmann 1988), é uma avaliação não-paramétrica que não requer suposições como normalidade e não-correlação serial entre os erros das técnicas consideradas. A estatística *SIGN* é obtida da seguinte forma:

$$SIGN = \frac{(\psi(N) - \frac{N}{2})}{2(N)^{\frac{1}{2}}} \quad (25)$$

¹⁷ Nesse caso, o erro de cada modelo é definido pela diferença entre o preço de mercado dos títulos considerados e suas respectivas estimativas, de acordo com a técnica utilizada.

em que o termo ψ_t denota o número de vezes em que, neste caso, os erros do modelo de MQNL superaram os erros derivados da técnica de evolução diferencial.

A hipótese nula é $H_0 : \psi(N) = (\frac{N}{2})$, contra a hipótese alternativa de que $H_1 : \psi(N) > (\frac{N}{2})$. Assume-se que N é o número de observações previstas por ambos os modelos e, se os erros de previsão da técnica tradicional de otimização são da mesma proporção que os obtidos pelos algoritmo de evolução diferencial em exatamente $(\frac{N}{2})$ observações, de acordo com o teste *SIGN*, a acurácia dos modelos pode ser considerada equivalente, sem diferença estatisticamente significativa. A estatística para o teste *SIGN* tem distribuição Gaussiana padrão. Assim, se a estatística *SIGN* é significativamente elevada, a hipótese nula de equivalência de predição pode ser rejeitada em favor da hipótese alternativa.

A Tabela 7 reporta os resultados dos testes estatísticos *MGN* e *SIGN* para avaliar os modelos de mínimos quadrados não-lineares e de evolução diferencial na estimação dos parâmetros das curvas de Nelson-Siegel e Svensson.

Tab. 7: Testes *MGN* e *SIGN* para avaliação das técnicas de MQNL e de ED na estimação dos parâmetros das funções de NSS.

Função de Nelson-Siegel		
Métodos	<i>MGN</i>	<i>SIGN</i>
MQNL Vs. ED	40,6703 (0,0000)	-1,3401 (0,0901**)
Função de Svensson		
Métodos	<i>MGN</i>	<i>SIGN</i>
MQNL Vs. ED	54,078 (0,0000)	6,5289 (0,0000)

(*) e (**) indicam significância a 1% e 5%, respectivamente.

Para a estimação dos parâmetros que definem a função de Nelson-Siegel, o teste paramétrico *MGN* nos permite inferir que o algoritmo de evolução diferencial é superior, em termos de precisão, a técnica de MQNL. Entretanto, em níveis de confiança de 1% e 5%, segundo o teste *SIGN*, as duas técnicas podem ser consideradas equivalentes. Porém, ainda assim, o algoritmo de ED permitiu a obtenção de parâmetros mais estáveis, o que confere maior adequação às teorias das expectativas do mercado financeiro e monetário. Para a estimação dos fatores da função de Svensson, ambos os testes estatísticos considerados nos permite inferir a superioridade do algoritmo evolutivo em relação à técnica tradicional de otimização. Apesar da proximidade dos resultados obtidos pelas abordagens consideradas. As Figuras 6 e 7 mostram as diferenças que se verificam nas curvas de juros de acordo com as técnicas de estimação empregadas.

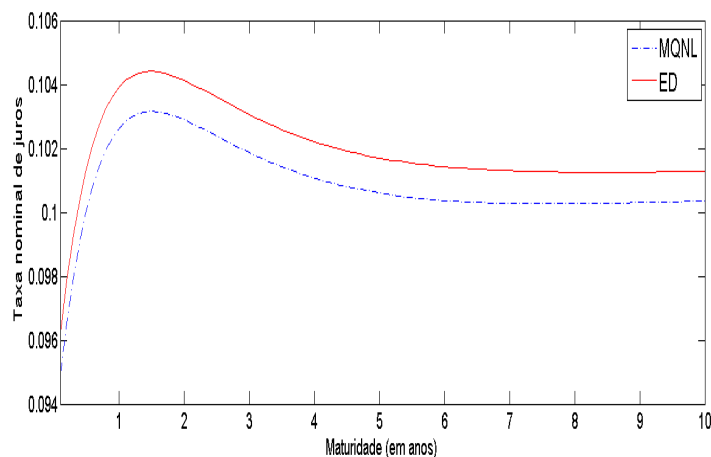


Fig. 6: Curva zero-cupom estimada para o dia 8 de Abril de 2008 com base na função de Nelson-Siegel de acordo com o método de MQNL e ED.

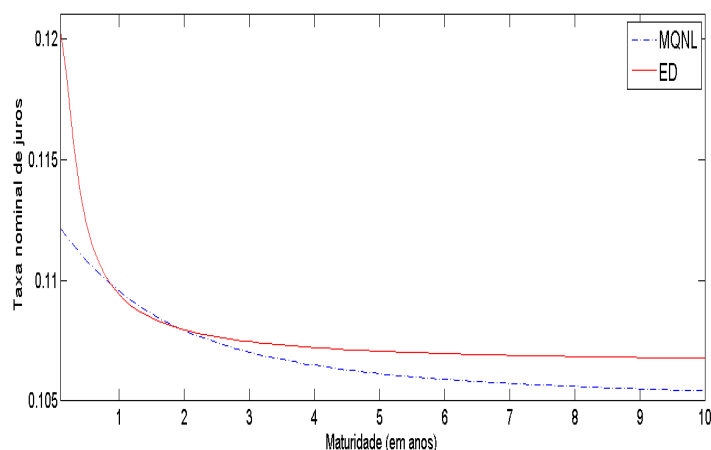


Fig. 7: Curva zero-cupom estimada para o dia 15 de Setembro de 2009 com base na função de Svensson de acordo com o método de MQNL e ED.

Nesse caso, como exemplo, selecionamos dois dias, 8 de Abril de 2008 e 15 de Setembro de 2009, para apresentar as curvas geradas a partir do modelo de Nelson-Siegel e Svensson, respectivamente, e que apresentaram formatos distintos. Pode-se notar que, apesar das diferenças entre as técnicas se mostrarem relativamente baixas, as curvas podem apresentar configurações distintas, em termos de nível, inclinação e curvatura.

Dessa forma, os resultados nos mostram que, investidores, *policy makers*, bancos centrais, e os demais participantes do mercado podem utilizar da técnica de evolução diferencial na estimação de modelos parcimoniosos da ETTJ de forma a obter uma curva mais robusta em relação ao mercado, assim como realizar predições mais acuradas sobre o futuro de taxas de juros, inflação e atividade econômica ao se considerar títulos indexados aos diversos índices macroeconômicos.

5 Conclusão

A estimação da estrutura a termo das taxas de juros é de elevada importância em termos macroeconômicos e financeiros. Para macroeconomistas, em um contexto de políticas monetárias, as taxas de juros futuras são indicadores essenciais das expectativas dos agentes quanto ao futuro das taxas de juros, inflação e atividade econômica. Em finanças, administradores de portfólios de renda fixa utilizam a curva de juros para marcação à mercado, enquanto gerenciadores de riscos a utilizam para precificar derivativos financeiros e realizar operações de *hedge*.

Neste contexto, as fórmulas funcionais de (Nelson e Siegel 1987) e (Svensson 1994) são as mais utilizadas por agentes de mercado, *policy makers*, pesquisadores e principais bancos centrais para a estimação da curva de juros zero-cupom. Entretanto, como não se verifica títulos públicos para todas as maturidades, é necessário estimar a curva de rendimentos via um processo de otimização, em geral não-linear. Dentre as técnicas mais utilizadas, o algoritmo de mínimos quadrados não-lineares (MQNL) se destaca. Porém, os métodos tradicionais de otimização implicam em algumas limitações, como o risco de falsa convergência e obtenção de parâmetros sensíveis às condições iniciais empregadas.

Este artigo propôs a utilização de um algoritmo de Evolução Diferencial (ED) para a estimação dos parâmetros que descrevem as curvas de Nelson-Siegel-Svensson. Essa técnica possui a vantagem de não depender das condições iniciais e diminui o risco da falsa convergência, uma vez que todos os pontos do espaço de busca apresentam uma probabilidade maior que zero de serem avaliados. Considerando como base de dados as Letras do Tesouro Nacional, estimou-se as curvas de NSS no mercado brasileiro por meio de um algoritmo de ED e com base na técnica de MQNL, para o período de Janeiro de 2007 a Dezembro de 2009. Os métodos foram comparados em termos de acurácia, estabilidade dos parâmetros e significância estatística.

Foram avaliados diferentes modelos de algoritmo de evolução diferencial, específicos quanto ao operador de mutação. Dentre eles, o algoritmo de ED que apresentou maior adequação à aplicação proposta foi aquele que realiza a operação de mutação de cada indivíduo em direção ao melhor indivíduo da população de cada geração. Os resultados mostraram que o algoritmo de evolução diferencial proposto resultou em curvas de juros mais precisas, ao se comparar com a metodologia de mínimos quadrados não-lineares. Além disso, o método de computação evolutiva gerou parâmetros mais estáveis, uma vez que a teoria monetária afirma que como os parâmetros revelam expectativas de curto e longo prazos, essas não podem se alterar em prazos curtos, o que confirma a adequação da metodologia proposta na estimação da ETTJ. Por fim, pode-se afirmar que, de acordo com os testes estatísticos aplicados, que o algoritmo de evolução diferencial superou a técnica de MQNL na estimação dos parâmetros das funções de NSS.

Os resultados foram promissores, de tal forma que essa técnica pode ser utilizada por agentes de mercado em geral e bancos centrais na obtenção da curva de juros, uma vez que permite a manutenção do sentido econômico dos parâmetros e adequação à teoria financeira por trás da estimação da estrutura a termo, assim como fornece estimativas mais acuradas de acordo com dados de mercado. Trabalhos futuros devem considerar a avaliação da técnica proposta em processos de decisão e em diferentes mercados, incluindo economias desenvolvidas e emergentes, assim como considerar e comparar com outros métodos baseados em computação evolutiva como o caso das Estratégias Evolutivas.

Referências

- Ang, A., Piazzesi, M. e Wei, M. (2006). What does the yield curve tell us about gdp growth?, *Journal of Econometrics* **131**(1–2): 359–403.
- BIS (2005). Zero-coupon yield curves: Technical documentation, *Bank for International Settlements Papers* (25).
- Bolder, D. e Stréliski, D. (1999). Yield curve modelling at the bank of canada, *Bank Of Canada Technical Report* (84): 1–16.
- Chambers, D. R., Charlenton, W. T. e Waldman, D. W. (1984). A new approach to estimation of the term structure of interest rates, *Journal of Financial and Quantitative Analysis* **19**(3): 233–252.
- Choudhry, M. (2004). *Analysing and interpreting the yield curve*, first edn, John Wiley and Sons, Singapore.
- Csajbok, A. (1998). Zero-coupon yield curve estimation from a central bank perspective, *National Bank of Hungary Working Paper* (2).
- Diament, P. (1993). Semi-empirical smooth fit to the treasury yield curve, *The Journal of Fixed Income* **1**: 55–70.
- Diebold, F. X. e Li, C. (2006). Forecasting the term structure of government bond yields, *Journal of Econometrics* **130**: 337–364.
- Diebold, F. X., Li, C. e Yue, V. Z. (2008). Global yield curve dynamics and interactions: A dynamic Nelson–Siegel approach, *Journal of Econometrics* **146**: 351–363.
- Fernández-Rodríguez, F. (2006). Interest rate term structure modelling using free knot splines, *The Journal of Business* **79**: 3083–3099.
- Gilli, M., Maringer, D. G. e Winker, P. (2008). *Handbook on Information Technology in Finance*, Springer-Verlag, Berlin, Heidelberg, chapter Applications of heuristics in finance.
- Gilli, M. e Winker, P. (2008). A review of heuristic optimization methods in econometrics, *Swiss Institute Research paper series* **1**(8): 8–12.
- Gimeno, R. e Nave, J. M. (2009). A genetic algorithm estimation of the term structure of interest rates, *Computational Statistics and Data Analysis* **53**: 2236–2250.

- Granger, J. W. C. e Newbold, P. (1977). *Forecasting economic time series*, New York: Academic Press.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*, University of Michigan Press.
- Ioannides, M. (2003). A comparison of yield curve estimation techniques using UK data, *Journal of Banking and Finance* **27**: 1–26.
- Krink, T., Mittnik, S. e Paterlini, S. (2009). Differential evolution and combinatorial search for constrained index-tracking, *Annals of Operational Research* **172**(2): 153–176.
- Krink, T. e Paterlini, S. (2011). Multiobjective optimization using differential evolution for real-world portfolio optimization, *Computational Management Science* **8**(1): 157–179.
- Lampinen, J. (2009). A bibliography of differential evolution algorithm, *Technical report*, <http://www2lutfi/jlampine/debibliohtm>.
- Lehmann, E. L. (1988). *Nonparametric statistical methods based on ranks*, New Jersey: Prentice Hall.
- Lin, L., Cao, L., Wang, J. e Zhang, C. (2004). The applications of genetic algorithms in stock market data mining optimization, *University of Technology Sydney White Papers*.
- Manousopoulos, P. e Michalopoulos, M. (2005). A comparison of yield curve estimation methods: The greek case, *Journal of Financial Decision Making* **1**(1): 33–46.
- Manousopoulos, P. e Michalopoulos, M. (2009). Comparison of non-linear optimization algorithms for yield curve estimation, *European Journal of Operational Research* **192**: 594–602.
- Mansi, S. A. e Phillips, J. (2001). Modelling the term structure from on-the-run treasury yield curve, *Journal of Financial Research* **24**: 545–564.
- Maringer, D. G. e Meyer, M. (2008). Smooth transition autoregressive models: New approaches to the model selection problem, *Studies in Nonlinear Dynamics & Econometrics* **12**(1): 1–19.
- Maringer, D. G. e Oyewumi, O. (1997). Index tracking with constrained portfolios, *Intelligent Systems in Accounting, Finance & Management* **15**(1–2): 57–71.
- Mckee, T. E. e Lensberg, T. (2002). Genetic programming and rough sets: A hybrid approach to bankruptcy classification, *European Journal of Operational Research* **138**(2): 436–451.
- Mitchell, M. (1998). *An introduction to Genetic Algorithms*, MIT Press, Cambridge, MA.
- Nelson, C. R. e Siegel, A. F. (1987). Parsimonious modeling of yield curves, *Journal of Business* **60**(4): 473–489.
- Price, K. V., Storn, R. M. e Lampinen, J. A. (2006). *Differential Evolution: A Practical Approach to Global Optimization*, second edn, Springer – Verlag, Berlin, Germany.
- Roudier, F. (2007). *Portfolio optimization and genetic algorithms*, Master's thesis, Swiss Federal Institute of Technology.
- Storn, R. M. e Price, K. V. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization* **11**(4): 341–359.
- Strutz, T. (2010). *Data Fiting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*, Vieweg and Teubner Verlag.
- Svensson, L. E. O. (1994). Estimating and interpreting forward interest rates: Sweden 1992-94, *International Monetary Fund Working Paper* (114).