# 2
# Aspectos Técnicos

## 2.1
## Fatoração de Nelson e Siegel (1987) da E.T.T.J.

A fatoração de Nelson e Siegel (1987) é uma maneira parcimoniosa de aproximar a E.T.T.J. em um dado período do tempo como uma combinação linear de três fatores. Seguindo a forma adotada em Diebold e Li (2006), em um período $t$, a taxa do título de maturidade $\tau$ é dada por:

$$Y_t(\tau) = \beta_{1,t} + \beta_{2,t}.\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t.\tau}\right) - \beta_{3,t}.\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t.\tau} - e^{-\lambda_t \tau}\right) + \varepsilon_t(\tau),$$

onde $\varepsilon_t(\tau)$ é um ruído com média zero por hipótese.

O parâmetro $\beta_{1,t}$ afeta igualmente as taxas de todas as maturidades, por tanto, pode ser interpretado como o nível da E.T.T.J.. Os choques econômicos capazes de afetar taxas de todas as maturidades indiscriminadamente são aqueles muito persistentes, portanto, pode-se dizer que $\beta_{1,t}$ é um fator de longo prazo.

O parâmetro $\beta_{2,t}$ tem peso inversamente proporcional à maturidade e está relacionado à inclinação da curva de juros. Por afetar mais as taxas curtas, esta fator está relacionado aos choques de transitórios da economia. Finalmente, o parâmetro $\beta_{3,t}$ tem peso decrescente na medida em que as maturidades se afastam das maturidades médias e relaciona-se à curvatura de E.T.T.J.. Estes três parâmetros são chamados de cargas dos fatores.

Além dos betas, há o parâmetro $\lambda_t$, que determina o decaimento da curva e a posição do máximo do fator que multiplica $\beta_{3,t}$. A exemplo do que foi feito em Diebold e Li (2006), o valor de $\lambda_t$ foi fixado em 0.0609, de tal maneira que o fator que multiplica $\beta_{3,t}$ atinja seu valor máximo na maturidade mediana (30 meses). A fixação do valor de $\lambda_t$ possibilita que, para cada período, a estimação dos demais parâmetros seja feita por MQO. Em De Pooter (2007), há evidências de que formas mais sofisticadas de estimação de $\lambda_t$ não trazem melhorias relevantes.

## 2.2
## Teste de Poder Preditivo de Giacomini e White (2006)

Com o de fito comparar o poder de previsão fora da amostra de duas ou mais metodologias de previsão para uma determinada série, Giacomini e White (2006) desenvolveram um novo teste estatístico, o qual é apresentado de maneira resumida na seqüência.

Sejam $f_t(\theta_{1,t})$ e $g_t(\theta_{2,t})$ os preditores concorrentes $\tau$ passos à frente para uma variável $Y_{t+\tau}$, e uma função perda $L_{t+\tau}(\cdot)$. Podemos definir as hipóteses do teste como sendo:

$$H_0 : E[L_{t+\tau}(Y_{t+\tau}, f_t(\hat{\theta}_{1,t})) - L_{t+\tau}(Y_{t+\tau}, g_t(\hat{\theta}_{2,t})) \mid \Phi_t] = 0, \text{ contra}$$

$$H_1 : E[abs(L_{t+\tau}(Y_{t+\tau}, f_t(\hat{\theta}_{1,t})) - L_{t+\tau}(Y_{t+\tau}, g_t(\hat{\theta}_{2,t}))) \mid \Phi_t] \geq \delta > 0,$$

onde $\Phi_t$ é um conjunto de informação contido na filtração gerada pelos eventos ocorridos até o período *t* inclusive.

Uma interpretação possível para a hipótese nula é que sob ela, dado o conjunto de informação $\Phi_t$, não é possível inferir qual dos dois preditores tende a gerar uma previsão $\tau$ passos à frente para a variável $Y_{t+\tau}$ mais acurada, sob a métrica da função perda $L_{t+\tau}(\cdot)$. Já sob a hipótese alternativa, um dos preditores tende a ser mais preciso que outro.

Este teste apresenta algumas vantagens em relação aos pré-existentes, dentre os quais, Diebold e Mariano (1995) e West (1996). Uma delas é levar em consideração a incerteza na especificação do modelo e na estimação dos parâmetros. Isso faz com que o teste não se limite a comparar apenas os modelos, e sim da metodologia empregada na previsão, que engloba, além do modelo em si, o método de estimação, a janela de dados considerada, entre outros fatores.

Outro avanço apresentado neste teste é a possibilidade de comparar o poder preditivo condicional a um dado conjunto de informação. Vale ressaltar o fato de o que teste incondicional, que se presta a responder se algum modelo é superior ou outro na média, é um caso particular do teste condicional.

No presente trabalho, estaremos interessados em comparar o poder preditivo incondicional ($\Phi_t$ é vazio) um período à frente ($\tau = 1$). Assim sendo, serão apresentadas, a seguir, as etapas da realização do teste sob estas circunstâncias:

1. Seja $\{\Delta L_{t+1}\}_{t=1}^{T} = \{L_{t+1}(Y_{t+1}, f_t(\hat{\theta}_{1,t})) - L_{t+1}(Y_{t+1}, g_t(\hat{\theta}_{2,t}))\}_{t=1}^{T}$. Regride-se a série $\{\Delta L_{t+1}\}_{t=1}^{T}$ em um vetor constante e igual a um;
2. Toma-se o $R^2$ desta regressão;
3. Compara-se $R^2.T$ com uma distribuição qui-quadrado com um grau de liberdade.
4. Caso a hipótese nula seja rejeitada, será uma evidência em favor do modelo com menor média amostral para a função perda.

O código para este e outro casos, programado em Matlab, está disponível na página da autora Raffaella Giacomini (www.econ.ucla.edu/giacomin).

## 2.3
## Algoritmo Stepwise

O algoritmo *stepwise* é um método para seleção, dentre um universo de potenciais variáveis explicativas, uma combinação que deve compor o modelo a ser utilizado.

Esta técnica remonta à década de 1970. Algumas versões ligeiramente diferentes deste algoritmo podem ser encontradas na literatura. Em particular, a forma utilizada no presente trabalho pode ser encontrada em Burgess (1999).

Seja $\Omega$ o universo de potenciais variáveis explicativas do modelo de interesse. Sejam $M_1$ e $M_2$ dois subconjuntos de $\Omega$, que definem dois modelos. Os modelos definidos por $M_1$ e $M_2$ são ditos vizinhos se e somente se:

1. $M_1 \subset M_2$ e $\#M_2 - \#M_1 \leq 1$[2], ou;
2. $M_2 \subset M_1$ e $\#M_1 - \#M_2 \leq 1$.

Seja, agora, $f(\cdot)$ a função que se pretende minimizar, definida do conjunto das partes de $\Omega$ nos números reais (ex.: um critério de informação do modelo

---

[2] No presente trabalho, $\#M_i$ denota o número de elementos do conjunto $M_i, i = 1,2$.

definido pelo subconjunto de $\Omega$). Seja, também, o operador $H(\cdot)$, definido do conjunto das partes de $\Omega$ nele próprio, satisfazendo:

1. se $H(M_1) = M_2$, então $M_1$ e $M_2$ são vizinhos, e;
2. se $H(M_1) = M_2$ e $M_1$ e $M_3$ são vizinhos, então $f(M_2) \leq f(M_3)$.

Ou seja, o operador $H(\cdot)$ leva um determinado subconjunto de $\Omega$ no seu vizinho para o qual a função $f(\cdot)$ assume o menor valor.

Isto posto, pode descrever o algoritmo *stepwise* da seguinte maneira:

1. parte-se de um conjunto $M_0 \subset \Omega$;
2. substitui-se $M_0$ por $H(M_0)$;
3. repete-se o passo 2 enquanto $H(M_0) \neq M_0$.

Vale notar que este algoritmo não garante que o conjunto $M_0$ resultante é aquele dentre o conjunto das partes de $\Omega$ para o qual $f(\cdot)$ assume o menor valor. Garante, apenas, que $f(\cdot)$ assume um valor menor ou igual em $M_0$ do que em qualquer um de seus vizinhos, ou seja, $M_0$ é um mínimo local.

Porém, como o número de elementos do conjunto das partes de $\Omega$ é igual a $2^{\#\Omega}$, uma busca extensiva pode ser inviável do ponto de vista do custo computacional, o que torna o *stepwise* uma ferramenta valiosa.