

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
BACHARELADO EM MATEMÁTICA APLICADA E COMPUTACIONAL

***Nowcasting* de PIB sob a Ótica  
de Aprendizagem Estatística**

Murilo Getlinger Coelho

MONOGRAFIA APRESENTADA AO INSTITUTO DE  
MATEMÁTICA E ESTATÍSTICA COMO REQUISITO PARCIAL  
PARA OBTENÇÃO DE GRAU DE BACHAREL EM  
MATEMÁTICA APLICADA E COMPUTACIONAL

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Florencia Leonardi

São Paulo  
21 de Maio de 2021



# Resumo

Murilo Getlinger Coelho. *Nowcasting de PIB sob a Ótica de Aprendizagem Estatística*. Monografia (Bacharelado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021.

O objetivo deste trabalho é avaliar a eficiência de métodos de *Nowcast* para a série do Produto Interno Bruto (PIB) brasileiro em comparação com modelos tradicionais de séries temporais. No caso do presente estudo foi escolhido como modelo para base de comparação um AR(4). Nos últimos anos houve um grande desenvolvimento desse tipo de técnica por conta de avanços computacionais. Na presente proposta trata-se de reunir um conjunto grande de variáveis relativas à economia brasileira por meio de diversas categorias, como produção industrial, mercado de trabalho, consumo de energia, mercado financeiro, balança comercial, entre outros, e a partir delas conseguir antecipar a variação do PIB para o período atual. Diferentes técnicas de aprendizagem estatística foram aplicadas e a previsão final foi feita filtrando os métodos por meio de testes baseados no *Model Confidence Set* (MCS).

**Palavras-chave:** Produto Interno Bruto (PIB). *Nowcasting*. Previsão. *Bridge Equations*. Aprendizagem Estatística. Lasso. *Support Vector Machine*.



# Abstract

Murilo Getlinger Coelho. **GDP Nowcasting with Statistical Learning**. Capstone Project Report (Bachelor). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2021.

The aim of this work is to evaluate the efficiency of Nowcasting methods to the Brazilian Gross Domestic Product (GDP) and compare the results to traditional time series techniques. The benchmark will be the AR(4). In recent years new techniques have been proposed to study nowcasting and prediction of economic variables, mostly due to the advancements of computational power. The idea is to gather various economic indicators of the Brazilian economy from many different categories, like industrial production, job market, energy consumption, financial market, trade balance, and others, and anticipate the variation of the current period's GDP. Multiple statistical learning techniques have been used and the final prediction was based on results from the Model Confidence Set (MCS) test.

**Keywords:** Gross Domestic Product (GDP). Nowcasting. Forecasting. Bridge Equations. Statistical Learning. Lasso. Support Vector Machine.



# Lista de Abreviaturas

AR	Auto-Regressivo
FAC	Função de Autocorrelação
FACP	Função de Autocorrelação Parcial
IBGE	Instituto Brasileiro de Geografia e Estatística
IIS	Saturação de Indicadores de Impulso ( <i>Impulse Indicator Saturation</i> )
IME	Instituto de Matemática e Estatística
MTE	Ministério do Trabalho e Emprego
RF	Floresta Aleatória ( <i>Random Forest</i> )
SVM	Máquina de Vetores de Suporte ( <i>Support Vector Machine</i> )
USP	Universidade de São Paulo

## Lista de Figuras

3.1	Lasso e Ridge - Normas L1 e L2. . . . .	12
4.1	Série do PIB trimestral brasileiro. . . . .	16
4.2	Estacionariedade de algumas variáveis . . . . .	17
5.1	Erro dos métodos . . . . .	23
5.2	<i>Nowcast</i> para algumas datas (1) . . . . .	24
5.3	<i>Nowcast</i> para algumas datas (2) . . . . .	25
5.4	<i>Nowcast</i> para algumas datas (3) . . . . .	26

## Lista de Tabelas

5.1	Especificação das datas . . . . .	20
5.2	Resultado MCS . . . . .	21
5.3	Erros dos Métodos . . . . .	22



# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Revisão Bibliográfica</b>	<b>3</b>
<b>3</b>	<b>Modelagem Teórica</b>	<b>7</b>
3.1	<i>Nowcasting</i> . . . . .	7
3.2	Aplicação de <i>nowcasting</i> para o PIB . . . . .	9
3.3	Componentes Principais . . . . .	10
3.4	Metodologias . . . . .	11
3.4.1	Lasso . . . . .	11
3.4.2	<i>Adaptive Lasso</i> . . . . .	12
3.4.3	<i>Elastic Net</i> . . . . .	12
3.4.4	<i>Support Vector Machine</i> . . . . .	13
3.5	Model Confidence Set . . . . .	13
<b>4</b>	<b>Análise Descritiva dos Dados</b>	<b>15</b>
<b>5</b>	<b>Resultados</b>	<b>19</b>
5.1	Especificação das Datas . . . . .	19
5.2	Métodos . . . . .	20
5.3	Exemplo de <i>Nowcast</i> . . . . .	23
5.4	Sugestões para pesquisas . . . . .	26
<b>6</b>	<b>Conclusões</b>	<b>27</b>
	<b>Referências</b>	<b>29</b>



# Capítulo 1

## Introdução

Previsão sempre foi um dos guias para toda decisão econômica, desde aplicações governamentais para apontar objetivos macroeconômicos, até utilizações no mercado financeiro visando uma vantagem sobre os demais concorrentes. O *Nowcast*, por outro lado, tem como foco prever o valor da próxima divulgação de variáveis relativas a períodos passados ou ao presente. Sua aplicação em áreas econômicas é relativamente recente, mas os resultados têm sido promissores. Técnicas de *forecast* comum apresentam bom poder preditivo para situações de relativa estabilidade econômica, mas pecam em momentos de alta variabilidade ou quebras estruturais.

O grande diferencial de modelos de *Nowcast* é a possibilidade da adição de novas atualizações no modelo, e com isso modificar a previsão sempre que novas informações são disponibilizadas pelos órgãos competentes. Trabalhos anteriores já buscaram mostrar que a adição de novas publicações é bastante relevante para aumentar o poder preditivo.

Com isso em mente, o objetivo deste trabalho é aplicar as teorias recentes de *Nowcast* para prever o Produto Interno Bruto (PIB) trimestral corrente no Brasil, e checar se para um país em desenvolvimento, durante uma crise mundial, os resultados encontrados para países desenvolvidos ainda são aplicáveis. A utilização deste tipo de modelo em situações de extrema incerteza como a pandemia do covid-19 é essencial para qualquer tomador de decisão, cada dado publicado tem o potencial de alterar muito as estimativas para a variação do trimestre corrente.

A modelagem foi feita com uma variação do modelo proposto no artigo de Garcia, Medeiros e Varsoncelos (2017), com a utilização de diversas especificações de modelos e com uma seleção dos melhores modelos baseada em *model confidence sets* (MCS). A base de dados é composta por 280 variáveis mensais e quatro trimestrais, com fotos semanais

ao longo do segundo semestre de 2020 e primeiro trimestre de 2021.

O Capítulo 2 contempla a revisão bibliográfica de técnicas de *nowcasting*, e especificamente aplicações para o PIB; o Capítulo 3 apresenta o arcabouço teórico do *nowcasting*, assim como uma explicação sobre cada um dos modelos utilizados e da escolha por MCS; no Capítulo 4 temos a análise descritiva dos dados e a explicação da base de dados; o Capítulo 5 descreve a especificação dos modelos e a análise dos resultados; o Capítulo 6 apresenta as principais conclusões do trabalho.

## Capítulo 2

# Revisão Bibliográfica

Técnicas de *nowcasting* vêm sendo utilizadas há bastante tempo, principalmente na área da pesquisa climática. Já seu uso para a pesquisa econômica é bem mais recente, ainda mais se forem consideradas as pesquisas desta modalidade voltadas para a realidade brasileira. Diversos autores apresentaram as vantagens da utilização deste método de previsão, mas existem desafios que têm de ser abordados para que os resultados possam ser considerados robustos e efetivos. Existem formas já aplicadas para controlar os possíveis problemas; Castle, Hendry e Kitov, (2013) e Banbura, Giannone, Modugno e Reichlin (2013) apresentaram algumas questões neste sentido e formas para resolvê-las.

No uso desta metodologia, como qualquer nova informação pode ser relevante, o número de variáveis que devem ser incluídas pode ser quase incomensurável; como existem diversas fontes para estes dados a sincronicidade entre as publicações quase nunca é possível, e a informação mais recente não necessariamente é referente ao último período; além disso, ainda mais no caso de variáveis com frequência trimestrais, como o PIB, o problema de diferentes frequências tem que ser abordado e contemplado. Modelos de *nowcasting* para variáveis macroeconômicas devem tentar corrigir o conjunto destes problemas.

Castle, Hendry e Kitov, (2013) separaram em sete itens principais os desafios para o desenvolvimento de modelos de *nowcasting* aplicado a variáveis macroeconômicas:

- falta de dados – alguns dados podem não estar disponíveis no momento de fazer a previsão do dado agregado;
- períodos de divulgação distintos – há diferenças temporais entre a publicação de diferentes indicadores;
- variação na base de dados – algumas variáveis podem ter dados faltantes em alguns

períodos ou atualizados após a primeira publicação;

- erro de medida – há falta de precisão na obtenção dos dados preliminares;
- quebras estruturais – como em qualquer série temporal, quebras estruturais de média e variância têm de ser levadas em consideração;
- mais variáveis,  $N$ , do que observações,  $T$  – como a base no *nowcasting* requer atualizações de alta frequência, há interação com um número potencialmente grande de variáveis;
- divulgação de dados não síncrona – diferentes dados são reportados com defasagens distintas.

Os autores propõem soluções para estes sete problemas especificamente para técnicas de *nowcasting*, nas quais pode acabar sendo ainda mais complicado conciliar todas as questões. Eles ilustram cada um dos erros apresentados em uma tentativa de *nowcast* do crescimento do PIB do Reino Unido.

Em um trabalho mais recente Garcia, Medeiros e Vasconcelos (2017) apresentam uma metodologia para o *nowcast* de inflação no Brasil, utilizando diversas técnicas de aprendizagem estatística, com o uso de *shrinkage* para reduzir a dimensão da base de dados e em seguida aplicando, por exemplo, modelos baseados em *Random Forests* (RF), LASSO, *Support Vector Machine* (SVM), entre outros. Em seguida os autores utilizam uma metodologia desenvolvida por Hansen, Lunde e Nason (2011) denominada *Model Confidence Set*, na qual a ideia é selecionar os modelos com maior poder preditivo estatisticamente significante. Garcia, Medeiros e Vasconcelos encontram resultados promissores com essa metodologia para a previsão de inflação no Brasil e com uma base de dados bastante variada e, por conta disso, utilizaremos uma metodologia similar, mas para uma variável trimestral.

Outros trabalhos que utilizam técnicas de *nowcasting* para o PIB incluem Liu, Matherson e Romeu, (2011), que fizeram 5 modelos diferentes de *forecasting* e *nowcasting* do PIB de 10 países da América Latina. Os autores mostram a importância clara da utilização de informações de indicadores mensais para aumentar a eficácia das previsões. Como seria de se esperar, os modelos apresentaram resultados mais assertivos quando os dados estavam mais completos, utilizando dados dos três meses anteriores à publicação do indicador: o erro de previsão com dados dos três meses é menor do que com dados apenas do primeiro mês. Encontram ainda grande importância para variáveis explicativas de países estrangeiros e preços de commodities, ressaltando a relevância da economia globalizada (principalmente os EUA) para o PIB dos países analisados.

O trabalho de Kahogara, (2015) procura fazer um *nowcasting* para o PIB brasileiro, da

década de 80 até 2014, com a antiga metodologia de cálculo do PIB. A principal conclusão do trabalho é o fato dos modelos propostos performarem, no mínimo, de forma igual ou melhor do que o modelo base, o auto-regressivo (AR), sugerindo o ganho com a utilização deste tipo de modelo. Conclui ainda que variáveis binárias não foram relevantes para uma melhora na eficácia das previsões, por conta de especificidades dos dados brasileiros.

Trabalhos que utilizam *nowcasting* para previsões de variáveis macroeconômicas em outros países incluem (i) PIB da Turquia em Modugno, Soybilgens e Yazgan, (2016), utilizando 14 indicadores econômicos turcos para estimar o PIB trimestral e anual; (ii) PIB da China em Yiu e Chow, (2011), no qual os autores fazem uso de 189 indicadores para a economia chinesa; (iii) PIB da Noruega em Aastveit e Trovik, (2012) – os autores fazem uma análise e mostram que o índice de ações do país é o conjunto de informações mais significativo para a previsão do PIB trimestral, seguido por indicadores de mercado de trabalho e produção industrial.

O indicador de pesquisa diária do PIB realizada pelo Banco Central foi testado contra modelos clássicos de *nowcasting* no trabalho de Bragoli, Metelli e Modugno, (2015). Mostraram a relevância de dados com maior frequência, como questionários e pesquisas, para aumentar a eficácia das previsões do PIB. Além disso, os autores avaliaram uma gama de indicadores contendo a pesquisa diária do PIB como preditivos do PIB real utilizando *nowcast*, explorando ainda mais as variáveis de alta frequência. Os autores testam apenas algumas variáveis como relevantes, e adicionam dados de fontes internacionais de divulgação, como a *Bloomberg*. Concluem que os resultados são estatisticamente tão bons quanto modelos clássicos de *forecast*.

A previsão de outras variáveis econômicas também foi amplamente explorada, como a inflação anual da zona do Euro em Banbura, Giannone e Reichlin, (2010) e americana em Giannone, Reichlin e Small, (2005), os ciclos econômicos em D'Agostino, Giannone, Lenza e Modugno, (2016) e o crescimento econômico global (série do FMI) em Ferrara e Marsilli, (2019), no qual os autores utilizam uma base de dados de 392 indicadores mensais (financeiros, da economia real, pesquisas de confiança) para 37 países distintos.

Em modelos de *forecasting* tradicional há o trabalho de Ericsson, (2017) sobre diferentes técnicas de *forecasting* para a dívida do governo americano. O trabalho abrange os impactos e riscos dos erros de previsão, além da importância de ter acuracidade para este indicador. O autor defende a utilização da técnica IIS – *Impulse Indicator Saturation* para detectar quebras estruturais nas variáveis, para mitigar os erros de previsão em pontos de inversão do ciclo de negócios e em expansões econômicas. O IIS é um teste genérico de detecção de quebras para um número desconhecido de eventos, que podem ocorrer a qualquer momento e em tempos, magnitudes e duração desconhecidas. Assim como em Kahogara,

(2015), os resultados com o IIS foram promissores.

Temos na literatura diversas aplicações de técnicas de *nowcasting* para variáveis macroeconômicas. Os principais resultados demonstram a melhora de desempenho que este tipo de modelagem pode trazer, principalmente quando é considerado um período pequeno, e particularmente no *nowcast* do período atual. Além disso, estatisticamente os modelos são tão consistentes quanto modelos tradicionais de *forecasting*, ou seja, a robustez matemática não é um problema. Por último, observa-se em quase todos os trabalhos que a assertividade da previsão aumenta consideravelmente no final dos períodos em questão, demonstrando o quão essencial é a atualização com as novas informações o quanto antes.



## Capítulo 3

# Modelagem Teórica

### 3.1 *Nowcasting*

A modelagem clássica de modelos de *nowcasting* utiliza, majoritariamente, a solução de *bridge equations*, um conjunto relativamente simplificado de equações por meio dos quais o *nowcast* é obtido com uma regressão da variável objetivo trimestral contra suas defasagens e alguns indicadores mensais. Nesta situação é necessário o preenchimento, por meio de interpolação, de algumas informações mensais não publicadas para um dado período. Existem casos na literatura de abordagens mais complexas para este tipo de modelagem, como o trabalho de Ferrara, Guegan e Rakotomalahy, (2010), que utiliza modelagem não paramétrica, com base em vizinhanças próximas, para prever os índices mensais necessários para a construção de um modelo paramétrico de previsão do PIB. Modelos nesta linha utilizam *bridge equations* que se mostram mais robustos tanto para *forecasts* quanto para *nowcasts*.

A modelagem básica utilizada neste trabalho é baseada em outra abordagem, similar à apresentada em Castle, Hendry e Kitov, (2013) e Banbura, Giannone e Reichlin, (2010). O objetivo é estimar uma variável com frequência trimestral (como o PIB) utilizando como insumo variáveis com maior frequência. Formalmente podemos escrever:  $\Omega_v$  é o conjunto de todas as informações disponíveis até o tempo  $v$ ;  $y_t^Q$  é o crescimento do PIB no período  $t$ . O problema do *nowcast* passa a ser então encontrar a projeção ortogonal de  $y_t^Q$  no espaço gerado por  $\Omega_v$ , ou seja:

$$\hat{y}_t^Q | \Omega_v = c + \sum_{i=1}^n \sum_{t_i=1}^{T_{i,v}} \beta_i x_{i,t_i} + \sum_{k=1}^{T_{Q,v}} \phi_k y_{t-k}^Q + \epsilon_t^Q$$

em que os  $\beta_i$  são os coeficientes das variáveis explicativas,  $\phi_k$  os coeficientes das defasagens

da própria variável explicada,  $\epsilon_t^Q$  é o erro,  $T_{i,v}$  o índice da última atualização da variável  $i$  no período  $v$  e  $T_{Q,v}$  o índice da última atualização da variável explicada no período  $v$ .

Sabe-se que o *nowcast* tem como uma de suas características principais a atualização da projeção conforme os dados são atualizados, formando assim uma sequência de projeções no tempo. É usual começar a projeção com quase nenhuma informação sobre o trimestre e o número de variáveis e observações incluídas crescer com o tempo. Tem-se, então, a sequência:

$$\hat{y}_t^Q | \Omega_v, \hat{y}_t^Q | \Omega_{v+1}, \hat{y}_t^Q | \Omega_{v+2}, \dots$$

onde denotamos  $v$ ,  $v+1$ ,  $v+2$  como datas consecutivas de novas informações disponíveis. Normalmente a distância entre os períodos acaba por ser bastante pequena devido ao número grande de variáveis.

Estudos já citados, como Giannone, Reichlin e Small, (2006), discutem o porquê da possível importância de toda e qualquer nova publicação de informações por órgãos competentes (como produção industrial, produção agrícola, emprego formal, entre outros) para a previsão do PIB. Por conta disso, cada nova informação exige uma atualização das projeções e revisão no modelo. A utilização, o quanto antes, destas novidades é, portanto, vista como extremamente relevante.

O período  $v+1$  comporta novos dados de algumas variáveis  $\{x_{j,T_{j,v+1}}, j \in J_{v+1}\}$ , e este novo conjunto é diferente do anterior por um dos seguintes motivos: ele contém novidades (atualizações de variáveis já existentes) e pode trazer atualizações/revisões de dados anteriores. Institutos de estatística brasileiros (IBGE, MTE, entre outros) têm como costume reavaliar dados anteriormente publicados.

A ideia passa a ser então: 1. Estimar uma previsão para o trimestre atual, condicionado ao conjunto de informações disponíveis; 2. Conforme novos dados são disponibilizados, atualizar a base; 3. Avaliar as novas informações na base de dados; 4. Avaliar fluxos incertos de informações futuras; 5. Avaliar o impacto das novas publicações de informações.

Assumindo normalidade dos dados, podemos reescrever a última equação em função de uma nova atualização de informações, e encontrar os coeficientes relativos às variáveis que foram atualizadas,  $b_{j,t,v+1}$ , de forma que:

$$\hat{y}_t^Q | \Omega_{v+1} = \hat{y}_t^Q | \Omega_v + \sum_{j \in J_{v+1}} b_{j,t,v+1} (x_{j,T_{j,v+1}} - E[x_{j,T_{j,v+1}} | \Omega_v])$$

Denotamos  $\hat{y}_t^Q | \Omega_{v+1}$  como o novo *forecast*,  $\hat{y}_t^Q | \Omega_v$  como o *forecast* anterior e  $(x_{j,T_{j,v+1}} -$

$\mathbb{E}[x_{j,T_{j,v+1}}|\Omega_v])$  a novidade advinda do último período. Podemos, então, escrever a mudança na previsão como a soma de todo o impacto das novas informações publicadas:

$$\hat{y}_t^Q|\Omega_{v+1} - \hat{y}_t^Q|\Omega_v = \sum_{j \in J_{v+1}} b_{j,t,v+1}(x_{j,T_{j,v+1}} - \mathbb{E}[x_{j,T_{j,v+1}}|\Omega_v])$$

onde  $J_{v+1}$  é o conjunto dos índices das variáveis que sofreram atualizações e  $b_{j,t,v+1}$  são os coeficientes dessas variáveis.

## 3.2 Aplicação de *nowcasting* para o PIB

Metodologias complexas de estimação de modelos de *nowcast* como modelos fatoriais ou *mixed data sampling* poderiam ser utilizados. Certamente são maneiras parcimoniosas de encontrar combinações em conjuntos extensos de dados com frequências mistas e com presença de assincronicidade dos dados. Contudo, os resultados finais obtidos com este tipo de modelo são de difícil interpretação. Por conta disso, neste trabalho será utilizada a metodologia, já brevemente descrita, de *bridge equations*, que buscam relacionar o PIB trimestral a variáveis mensais agregadas de uma ou mais séries - no nosso caso utilizaremos ainda como variáveis explicativas os componentes principais das bases de dados mensais. Castle, Hendry e Kitov citam em seu trabalho de 2013 que modelos baseados neste método são amplamente utilizados por Bancos Centrais de diversos países, por conta de sua fácil interpretação e, portanto, aplicabilidade em políticas públicas. O objetivo passa a ser estimar uma equação para a variação do PIB,  $\Delta y_{t_q}$  em uma frequência trimestral:

$$\Delta \hat{y}_{t_q} = c + \sum_{i=1}^p \alpha_i \Delta y_{t_{q-1}} + \sum_{j=0}^p \sum_{i=1}^k \beta_{i,j} z_{i,t_q-j} + u_{t_q}$$

para  $t = 1, 2, \dots, T$ , sendo que os  $k$  indicadores trimestrais  $z_{i,t_q}$  são transformados em estacionários,  $p$  é o *lag* e assumindo que  $u_t$  são resíduos IID. Observa-se que temos  $p$  *lags* para a variação do PIB (coeficientes  $\alpha_i$ ) e  $p$  *lags* para os indicadores explicativos, além do contemporâneo, se disponível (coeficientes  $\beta_{i,j}$ ).

O método usado para comparação de performance será um AR(4) baseado na série sem ajuste para sazonalidade do crescimento trimestral do PIB, disponibilizada pelo IBGE:

$$\hat{y}_t^Q = c + \sum_{i=1}^4 \beta_i y_{t-i}^Q + \epsilon_t^Q$$

com  $c$  constante e  $\epsilon_t^Q$  o ruído branco trimestral, de tal forma que  $\epsilon_t^Q \sim N(0, \sigma_\epsilon^2)$ .

### 3.3 Componentes Principais

A utilização dos componentes principais neste trabalho é certamente essencial para conseguirmos extrair toda a variação relevante de uma base de dados com muito mais variáveis do que observações no tempo. Assim sendo, utilizaremos sempre a transformação da base para todas as variáveis mensais.

A ideia dos componentes principais é projetar toda a base de dados em um novo sistema de coordenadas com a variância total sendo decrescente nas coordenadas, ou seja, a primeira coordenada (que corresponde ao primeiro componente) será a com maior variância, em seguida a segunda coordenada, e assim por diante.

Seja  $\mathbf{X}$  uma matriz  $n \times p$ , com média zero para cada uma das colunas.  $p$  representa o número de indicadores que temos na base e  $n$  o número de observações. A transformação efetuada pelas componentes principais é definida por um conjunto de tamanho  $l$  de vetores de peso com dimensão  $p$ ,  $\mathbf{w}^{(k)} = (w_1, \dots, w_p)^{(k)}$  que basicamente leva cada linha  $\mathbf{x}_{(i)}$  de  $\mathbf{X}$  para um novo vetor já transformado  $\mathbf{t}_{(i)} = (t_1, \dots, t_l)_{(i)}$ :

$$t_k(i) = \mathbf{x}_{(i)} \cdot \mathbf{w}^{(k)} \quad \text{for} \quad i = 1, \dots, n \quad k = 1, \dots, l$$

com  $l \leq p$ . Claro que o que faz sentido será escolher sempre  $l < p$  para que consigamos diminuir a dimensionalidade da base de dados.

Para maximizar a variância teremos o primeiro vetor de pesos  $\mathbf{w}_{(1)}$  de tal forma que:

$$\mathbf{w}_{(1)} = \arg \max_{\|\mathbf{w}\|=1} \left\{ \sum_i (t_1)_{(i)}^2 \right\} = \arg \max_{\|\mathbf{w}\|=1} \left\{ \sum_i (\mathbf{x}_{(i)} \cdot \mathbf{w})^2 \right\}$$

Em forma matricial teremos:

$$\mathbf{w}_{(1)} = \arg \max_{\|\mathbf{w}\|=1} \{ \|\mathbf{X}\mathbf{w}\|^2 \} = \arg \max_{\|\mathbf{w}\|=1} \{ \mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w} \}$$

Encontrando a primeira componente principal ( $t_{1(i)} = \mathbf{x}_{(i)} \cdot \mathbf{w}_{(1)}$ ) conseguimos descobrir os componentes subsequentes. O  $k$ -ésimo componente será obtido retirando de  $\mathbf{X}$  os  $k-1$  componentes já calculados:

$$\hat{\mathbf{X}}_k = \mathbf{X} - \sum_{s=1}^{k-1} \mathbf{X}\mathbf{w}_{(s)}\mathbf{w}_{(s)}^T$$

e, em seguida, calculando o vetor que maximiza a variância dessa matriz resultante:

$$\mathbf{w}_{(k)} = \arg \max_{\|\mathbf{w}\|=1} \left\{ \|\hat{\mathbf{X}}_k \mathbf{w}\|^2 \right\} = \arg \max \left\{ \frac{\mathbf{w}^T \hat{\mathbf{X}}_k^T \hat{\mathbf{X}}_k \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \right\}$$

No fim teremos como vetores de peso os autovetores de  $\mathbf{X}^T \mathbf{X}$ .

## 3.4 Metodologias

Nesta sessão serão discutidas quais técnicas foram utilizadas para prever a variação do PIB trimestral a partir da base de dados transformada com componentes principais.

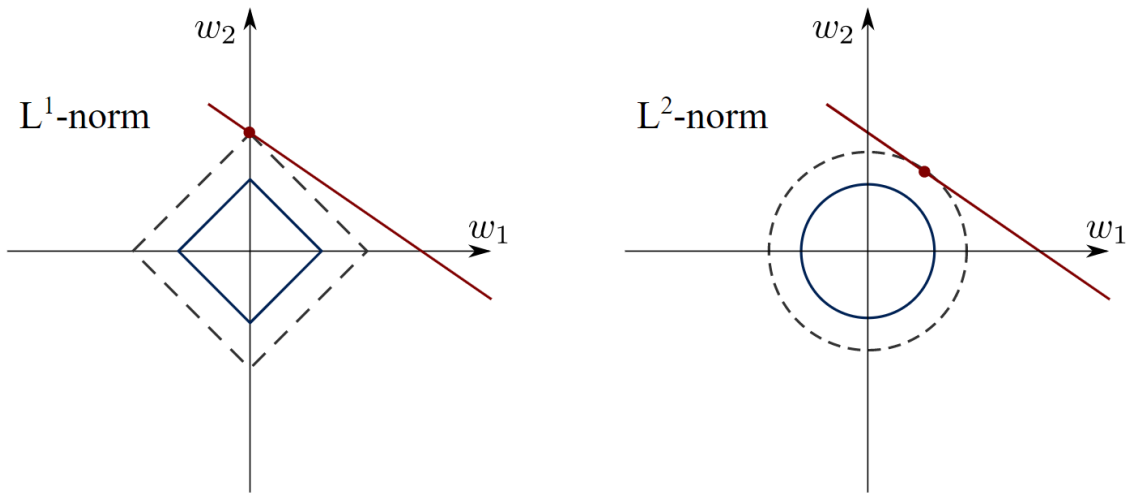
### 3.4.1 Lasso

Seja  $N$  o número de observações em uma base de dados, e cada observação composta de  $p$  variáveis. Considere  $y_i$  como a variável objetivo (no nosso caso a variação do PIB trimestral) e  $\mathbf{x}_i := (x_1, \dots, x_p)^T$  o vetor de covariadas para a observação  $i$ . O objetivo do lasso é encontrar  $\beta_0$  e  $\beta$  satisfazendo:

$$\min_{\beta_0, \beta} \left\{ \sum_{i=1}^N (y_i - \beta_0 - \mathbf{x}_i^T \beta)^2 \right\} \text{ sujeito a } \sum_{j=1}^p |\beta_j| \leq t$$

onde  $t$  é um hiperparâmetro que define o grau de regularização.

Teremos com isso, que os coeficientes das variáveis menos relevantes para explicar a variância de  $y$  serão todos levados a zero, deixando com menos variáveis relevantes na regressão. Geometricamente podemos entender essa técnica como encontrar o ponto tangente no primeiro quadrante entre o segmento de reta vermelho e a bola de norma 1, como na Figura 3.1. Uma outra técnica possível, mas que penaliza menos por usar a norma 2, seria a *ridge regression*, à direita na Figura 3.1.



Fonte: Wikipedia

**Figura 3.1:** Lasso e Ridge - Normas L1 e L2.

### 3.4.2 Adaptive Lasso

A ideia do lasso adaptativo é praticamente igual à ideia do lasso, apenas adicionando um termo na função de penalização para corrigir por problemas de viés que muitas vezes ocorrem com o lasso. A ideia é encontrar  $\beta_0$  e  $\beta$  satisfazendo:

$$\min_{\beta_0, \beta} \left\{ \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 \right\} \text{ sujeito a } \sum_{j=1}^p w_j |\beta_j| \leq t$$

onde  $t$  é um hiperparâmetro que define o grau de regularização.

Normalmente os pesos  $w_j$  são escolhidos a partir de uma estimativa já previamente feita dos betas, de tal forma que fique realmente uma adaptação baseada no lasso.

### 3.4.3 Elastic Net

A técnica do *elastic net* é uma generalização do lasso e do *ridge regression*, basicamente tomando uma combinação convexa das penalidades das duas técnicas e utilizando em conjunto. Aqui, o  $\beta$  escolhido será de tal forma que:

$$\hat{\beta} \equiv \underset{\beta}{\operatorname{argmin}} (\|y - X\beta\|^2 + \lambda_2 \|\beta\|^2 + \lambda_1 \|\beta\|_1)$$

onde  $\lambda_2$  e  $\lambda_1$  são hiperparâmetros que devem ser parametrizados.

### 3.4.4 *Support Vector Machine*

A ideia do SVM é construir um conjunto de hiperplanos em um espaço com dimensão maior do que o da base e utilizar essa separação para classificação ou, no nosso caso, previsão. Gostaríamos, então, de encontrar o hiperplano com distância maior para os pontos de nossa amostra. No caso deste trabalho o objetivo é usar esse mesmo conjunto de técnicas para estimar uma regressão, então ao invés de encontrar um hiperplano para separar e classificar a amostra encontraremos uma função que dita a relação entre as variáveis explicativas e as variáveis explicadas. Utilizando a mesma nomenclatura para as variáveis que utilizada anteriormente teremos definidos implicitamente os coeficientes  $w$  de tal forma que:

$$\text{minimizar } \frac{1}{2} \|w\|^2$$

$$\text{sujeito a } |y_i - w \cdot x_i - b| \leq \varepsilon \text{ para todo } i$$

onde  $y_i$  é uma observação de nossa variável objetivo,  $x_i$  é um vetor com todas as variáveis explicativas para a observação  $i$ ,  $b$  é o intercepto e  $\varepsilon$  é um hiperparâmetro utilizado para definir a distância máxima que a previsão deve estar do valor real.

## 3.5 Model Confidence Set

Após rodar todos os métodos faremos uma comparação dos erros de previsão e retiraremos da média final os que tiverem desempenho pior para previsão. Faremos isso de acordo com o MCS, apresentado por Hansen, Lunde e Nason (2011). O teste funciona da seguinte maneira, seja  $l_{i,t}$  o erro de previsão do método  $i$  para o período  $t$ . Defina:

$$d_{ij,t} = l_{i,t} - l_{j,t}, \quad i, j = 1, \dots, m, \quad t = 1, \dots, n,$$

como a diferença entre o erro dos métodos.

O teste de hipótese utilizado é:

$$H_0 : \quad \mathbb{E}[d_{ij}] = 0 \quad \text{para todo } i, j = 1, 2, \dots, m$$

$$H_1 : \quad \mathbb{E}[d_{ij}] \neq 0 \quad \text{para algum } i, j = 1, 2, \dots, m$$

A estatística de teste construída a partir de *bootstrapping* é então:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{var}(\bar{d}_{ij})}} \quad \text{para } i, j \in M$$

Se, por acaso, um dos métodos tiver um erro de previsão muito maior que o dos outros esse teste irá reconhecer esse método, retirá-lo da amostra, e rodar o teste novamente, até que não exista uma diferença estatisticamente significativa entre os erros de previsão dos métodos remanescentes.



## Capítulo 4

# Análise Descritiva dos Dados

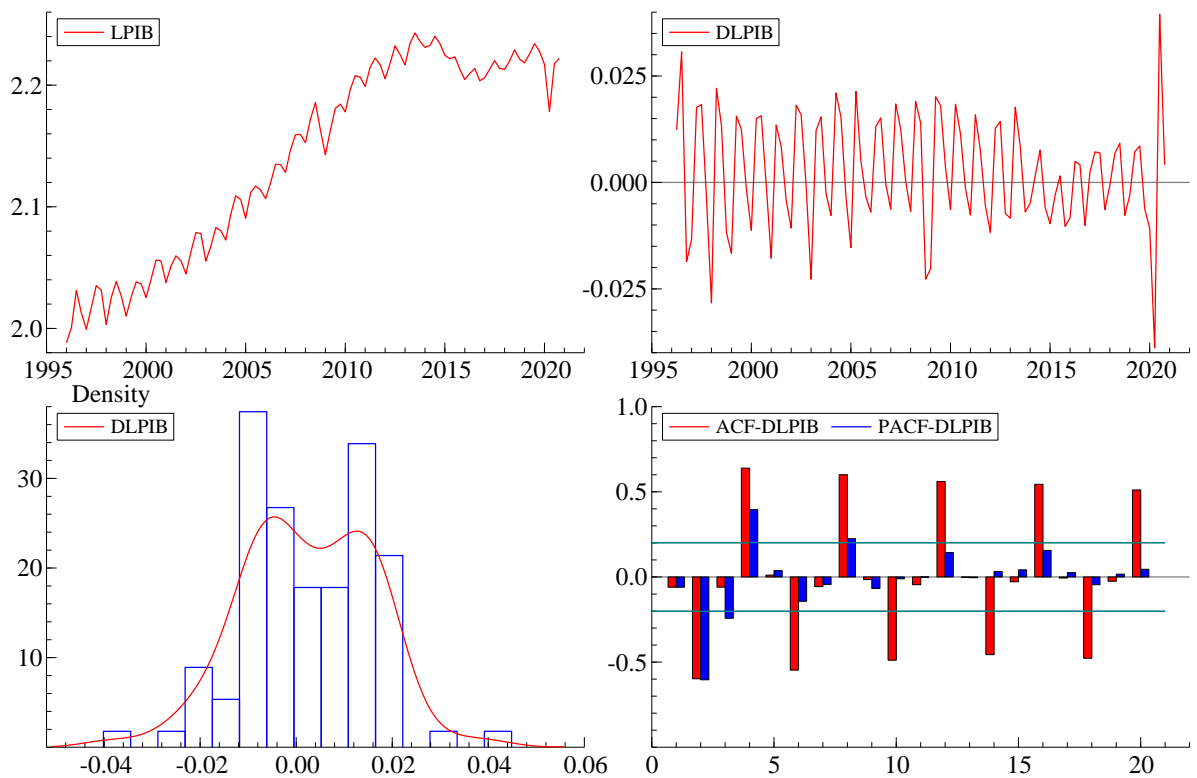
A base de dados deste trabalho consiste em um total de 280 indicadores mensais, considerando a economia brasileira, tanto no setor real (Produção Industrial, Produção de automóveis, Mercado de Trabalho, Consumo de Energia Elétrica), quanto no financeiro e commodities. Além disso, como já citado, existe grande importância em utilizar indicadores de economias estrangeiras para a previsão do PIB brasileiro, e, portanto, dados de Produção Industrial, Balança comercial e índices do mercado financeiro de países chave foram adicionados à base.

A janela escolhida para o estudo parte da série do índice do PIB a preços de mercado, disponibilizada pelo IBGE, após o novo ajuste de metodologia. Esta série está disponível, trimestralmente, a partir de 1996, tendo a média de 1995 como base (=100). Valls e Viera, (2014) em seu trabalho “Um estudo sobre os ciclos de negócios brasileiro” mostram detalhadamente o funcionamento dos ciclos de negócio brasileiro no período estudado e as diversas quebras estruturais que ocorreram na economia do país. Somente na década de 80 pode-se observar, além da crise interna no Brasil, com 24 trimestres de recessão, aumento nas taxas de juros e aumento nos preços internacionais do petróleo, que acabaram aumentando a situação das dívidas de países da América Latina. O processo de hiperinflação sofrido pelo Brasil após estas turbulências gerou diversas tentativas fracassadas de planos de contenção, e cada um deles causando quebras estruturais nas séries de dados.

Mesmo em períodos mais recentes vemos o mesmo ocorrer, o plano Collor em 1990, seguido por crises financeiras entre 1995 e 1999 trazem ainda mais dificuldades em lidar com séries econômicas referentes ao Brasil. Observa-se claramente o número enorme de quebras que as séries dos índices brasileiros carregam, e, portanto, a necessidade clara de uma preocupação constante com quebras estruturais em modelos de previsão do PIB.

Por conta disso, e da disponibilidade de algumas variáveis com frequência maior, a janela escolhida no fim começou no primeiro trimestre de 2010 (após os principais choques de 2008 já terem sido dissipados).

Temos na figura seguinte a série do PIB disponibilizada pelo IBGE. Primeiramente a série em logaritmo, seguida pela primeira diferença do logaritmo da série, sua distribuição, e a Função de Autocorrelação (FAC) e a Função de Autocorrelação Parcial (FACP) da série - essas últimas duas utilizadas para checar as defasagens da variável úteis para previsão. A memória é clara, e além do ciclo óbvio de um ano é possível observar pequenos ciclos a cada 2 períodos.



Fonte: Ipeadata

**Figura 4.1:** Série do PIB trimestral brasileiro.

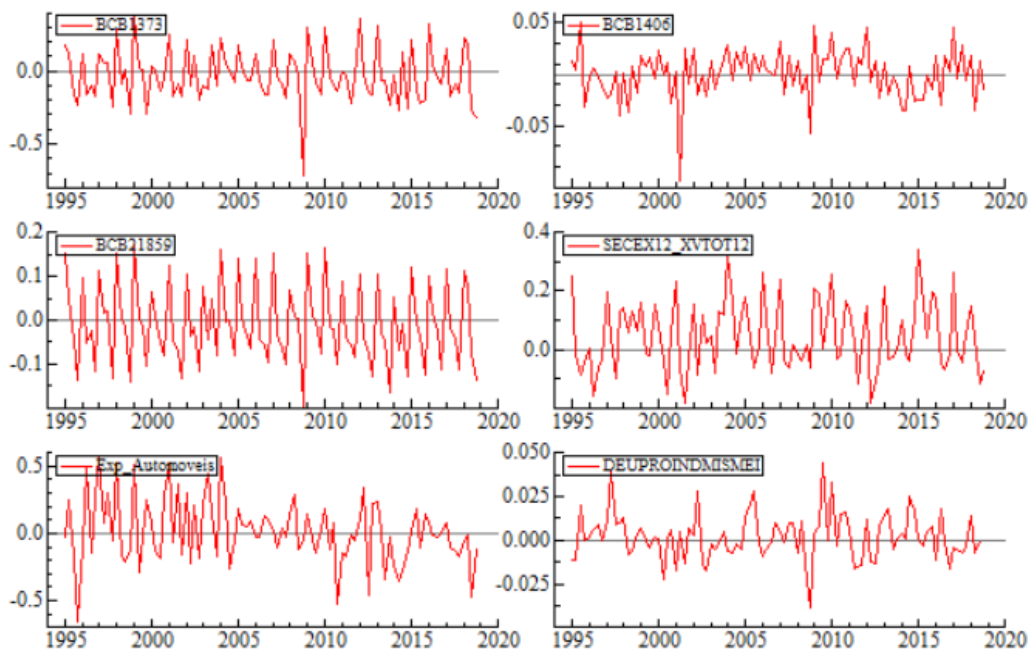
A utilização de diversas séries é de extrema importância, uma vez que, ao utilizar os componentes principais precisamos extrair realmente as dimensões com variação mais relevante para a previsão. Levando isso em consideração, a escolha das variáveis inclui diversos indicadores que podem ajudar condicionalmente a previsão do PIB.

Todas as variáveis têm periodicidade mensal, exceto o PIB, e o crescimento médio até o momento no trimestre do IBC-Br, da Produção Industrial, da Pesquisa Mensal do Comércio

e da Pesquisa Mensal de Serviços. Vale destacar novamente a presença de variáveis relativas a outros países, que, como já abordado, são essenciais para a previsão do PIB em um mundo globalizado. Outro ponto relevante foi a necessidade de encadear alguns índices, por conta de mudanças na metodologia de cálculo. Isso ocorreu com praticamente todas as séries de produção industrial divulgadas pelo IBGE.

Um problema já citado é a diferença entre as defasagens e o período à qual se referem cada uma das novas atualizações de variáveis. Por exemplo, a série do índice Bovespa é atualizada imediatamente após o acontecimento, ou seja, o índice de cada mês sai imediatamente após o fechamento da bolsa no último dia de cada mês, enquanto séries como a de produção industrial tem atualização referente à um período anterior. Como o tratamento foi feito com bases de dados a cada semana esse problema foi mitigado, com os *nowcasts* e projeções sempre representando a base de dados disponível na data de referência.

Castle, Hendry e Kitov, (2013) sugerem a utilização de variáveis sempre estacionárias. Por conta disso foram feitas transformações necessárias e testes de raízes unitárias nas variáveis para que as séries utilizadas como insumo fossem sempre  $I(0)$  trabalhando com a diferença dos logaritmos. Algumas variáveis após essas transformações podem ser vistas na Figura 4.2.



Fontes: IBGE, Ipeadata, FRED

**Figura 4.2:** Estacionariedade de algumas variáveis

As variáveis apresentadas são, respectivamente, Produção de Automóveis, Consumo Total de Energia Elétrica, Indicador Geral de Produção Industrial, Exportações FOB, Exportações de Automóveis e Produção Industrial Alemã.

# Capítulo 5

## Resultados

### 5.1 Especificação das Datas

Tendo em mente as diferenças em defasagens e periodicidades das divulgações de dados, algumas especificações de métodos foram escolhidos, uma para cada data base. A tabela a seguir resume as características de cada uma das datas:

As datas podem ser separados em três grupos:

- (a) Números 1 a 9: base de dados para esses modelos não inclui ainda a divulgação do PIB para o terceiro trimestre, então quanto estamos analisando a projeção para o quarto trimestre é, basicamente, a previsão para o segundo horizonte. Como as bases de dados já estão no quarto trimestre a projeção para o quarto trimestre é denominada *nowcasting*;
- (b) Números 10 e 11: aqui já temos a divulgação do PIB do terceiro trimestre, então as projeções para o quarto trimestre ocorrem como projeções para o primeiro horizonte. Segue sendo *nowcasting* pois as bases ainda são obtidas no quarto trimestre de 2020;
- (c) Números 12 a 19: assim como no último item já temos a divulgação do PIB do terceiro trimestre, a diferença aqui é que a data base já é em 2021. Isso ocorre pois o dado do PIB é divulgado aproximadamente dois meses depois do final do trimestre e, além disso, outros dados sobre o quarto trimestre ainda estavam sendo divulgados. Como a data base já é no primeiro trimestre de 2021 temos um *postcast* quanto estamos olhando para a previsão do quarto trimestre de 2020.

O objetivo de rodar para todas estas datas foi ter uma forma de comparar os erros conforme novas atualizações de dados são incorporados à base de dados. O arcabouço teórico

	Método	Número	Data Base
1	Nowcast	1	07/10/2020
2	Nowcast	2	14/10/2020
3	Nowcast	3	21/10/2020
4	Nowcast	4	28/10/2020
5	Nowcast	5	04/11/2020
6	Nowcast	6	11/11/2020
7	Nowcast	7	18/11/2020
8	Nowcast	8	25/11/2020
9	Nowcast	9	02/12/2020
10	Nowcast	10	09/12/2020
11	Nowcast	11	16/12/2020
12	Nowcast	12	22/12/2020
13	Postcast	13	06/01/2021
14	Postcast	14	13/01/2021
15	Postcast	15	20/01/2021
16	Postcast	16	27/01/2021
17	Postcast	17	04/02/2021
18	Postcast	18	10/02/2021
19	Postcast	19	17/02/2021
20	Postcast	20	24/02/2021

**Tabela 5.1:** *Especificação das datas*

e testes efetuados em outros estudos chegaram à conclusão que toda nova informação divulgada pelos órgãos relevantes tem impacto positivo para a melhora dos modelos. A ideia é, então, testar a aplicabilidade desta teoria para a nova metodologia do cálculo do PIB brasileiro.

## 5.2 Métodos

O resultado dos métodos pode ser resumido pelo gráfico abaixo. As linhas representam o erro médio absoluto para a previsão de um horizonte em diversos períodos do tempo. Seja  $i \in \{1, 2, \dots, 20\}$  o índice para as diferentes bases (com dados relativos apenas aos dados publicados até a data base, de acordo com a Tabela 5.1) e seja  $j \in \{1, 2, \dots, 9\}$  o índice para avaliar a janela móvel de tempo utilizada para prever - aqui estamos fazendo um *backtest* dos métodos para avaliar os erros (com  $j = 1$  teremos a previsão para o quarto trimestre de 2018, com uma base com as mesmas variáveis publicadas de acordo com a Tabela 5.1,  $j = 2$  a previsão para o primeiro trimestre de 2019, e assim por diante). O erro para o período  $i$  será calculado de acordo com a fórmula:

$$erro_i = \frac{1}{9} \sum_{j=1}^9 |\hat{y}_m^i - y^i|$$

onde  $\hat{y}_m^i$  é a previsão do método  $m$  com a data base  $i$  e  $y^i$  é o valor realizado.

Com essa série de erros para cada um dos métodos, a escolha dos métodos que seriam escolhidos para a média foi feita a partir do MCS. O resultado do MCS pode ser visto pela tabela 5.2.

	Data Base	Métodos Escolhidos
1	07/10/2020	Lasso, Ada. Lasso
2	14/10/2020	Lasso, Ada. Lasso, SVM
3	21/10/2020	Lasso, Ada. Lasso, SVM
4	28/10/2020	Lasso, Ada. Lasso
5	04/11/2020	Lasso, Ada. Lasso, SVM
6	11/11/2020	Lasso, Ada. Lasso, SVM
7	18/11/2020	Lasso, Ada. Lasso, SVM
8	25/11/2020	Lasso, Ada. Lasso, SVM
9	02/12/2020	Lasso, Ada. Lasso, SVM
10	09/12/2020	Lasso, Ada. Lasso, SVM
11	16/12/2020	Lasso, Ada. Lasso, Elastic Net
12	22/12/2020	Lasso, Ada. Lasso, Elastic Net
13	06/01/2021	Lasso, Ada. Lasso, Elastic Net
14	13/01/2021	Lasso, Ada. Lasso
15	20/01/2021	Lasso, Ada. Lasso
16	27/01/2021	Lasso, Ada. Lasso, Elastic Net
17	04/02/2021	Lasso, Ada. Lasso
18	10/02/2021	Lasso, Ada. Lasso, Elastic Net
19	17/02/2021	Lasso, Ada. Lasso
20	24/02/2021	Lasso, Ada. Lasso

**Tabela 5.2:** *Resultado MCS*

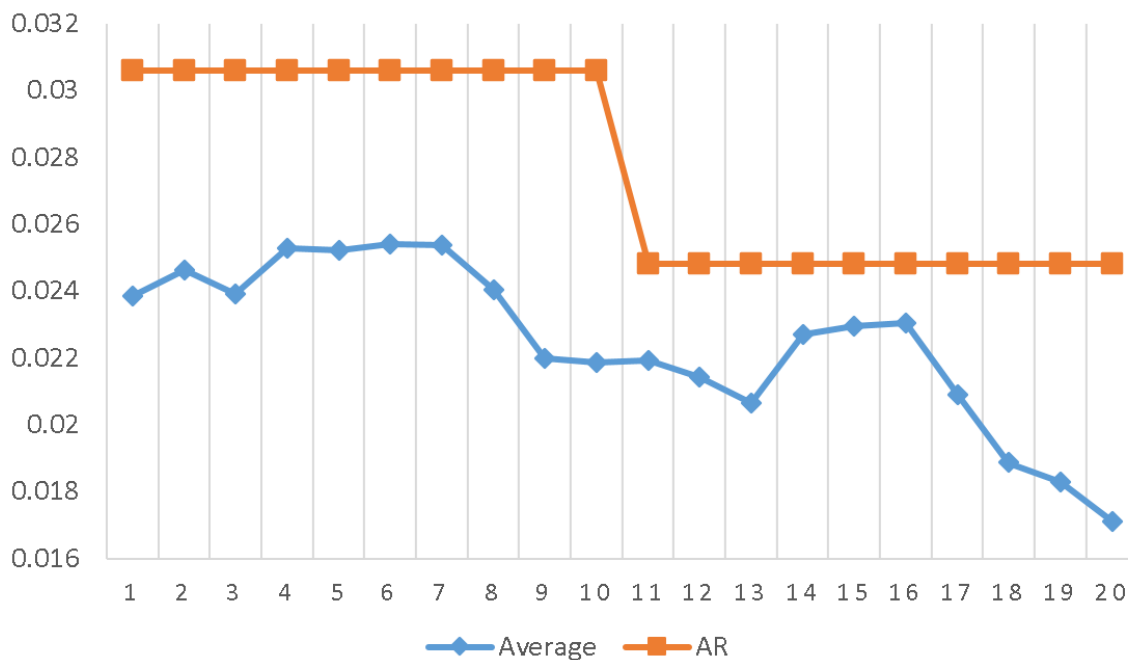
A Tabela 5.3 traz a comparação dos erros dos métodos para todas as datas base utilizadas, pode-se ver que até as últimas três datas os métodos performam de forma bastante similar, somente no final que há uma diferenciação significativa para a média dos métodos.

	Average	AR	Lasso	Adaptive Lasso	Elastic Net	SVM
1	0.02384	0.03061	0.023678	0.023966	0.02341	0.025539
2	0.024635	0.03061	0.024807	0.02473	0.025459	0.025285
3	0.023931	0.03061	0.02331	0.023682	0.02369	0.025043
4	0.025288	0.03061	0.023608	0.023907	0.023804	0.025408
5	0.025211	0.03061	0.025392	0.026052	0.025778	0.025226
6	0.025419	0.03061	0.025456	0.026045	0.025973	0.025224
7	0.025384	0.03061	0.025604	0.026565	0.025779	0.02542
8	0.024032	0.03061	0.025604	0.026565	0.025779	0.02542
9	0.021999	0.03061	0.025264	0.025866	0.0244	0.02545
10	0.021852	0.03061	0.025325	0.025933	0.024496	0.025446
11	0.021928	0.024809	0.02017	0.020961	0.019999	0.025752
12	0.021423	0.024809	0.020873	0.022028	0.021747	0.025776
13	0.020646	0.024809	0.020911	0.019661	0.01988	0.025626
14	0.022707	0.024809	0.021817	0.021398	0.023922	0.025403
15	0.022943	0.024809	0.021612	0.024151	0.023409	0.025362
16	0.023049	0.024809	0.023858	0.024393	0.022845	0.025364
17	0.020906	0.024809	0.02192	0.021799	0.023062	0.025446
18	0.018862	0.024809	0.021298	0.021143	0.020797	0.025446
19	0.018291	0.024809	0.020042	0.019742	0.020082	0.025303
20	0.017108	0.024809	0.020049	0.019747	0.020082	0.025305

**Tabela 5.3:** *Erros dos Métodos*

Pode-se observar que as informações obtidas pelas atualizações dos órgãos competentes, pelo menos pelo critério do erro médio, estão realmente melhorando a performance do método na previsão do crescimento do  $\log(\text{PIB})$ , principalmente no final do período, ao chegar mais perto de ter toda a base completa. O nosso método principal é a média dos métodos que tem poder de previsão estatisticamente igual pelo MCS, e esse é o que está representado no gráfico abaixo. Pode-se perceber que o método está performando melhor do que o modelo base de comparação (AR(4)) sem variáveis explicativas para todas as bases de dados.





**Figura 5.1:** Erro dos métodos

### 5.3 Exemplo de *Nowcast*

O resultado do *Nowcast* é prever o PIB no período corrente, e como diversos horizontes foram testados apenas um dos resultados será apresentado. Os gráficos a seguir (Figuras 5.2, 5.3 e 5.4) são as previsões de *nowcast* para o PIB do último trimestre de 2020 com os dados sendo atualizados a cada data base, de acordo com os números especificados anteriormente. É interessante notar como há variabilidade nas previsões conforme novas atualizações são incorporados na base de dados, ressaltando a importância da adição dessas novas informações divulgadas pelos órgãos competentes.

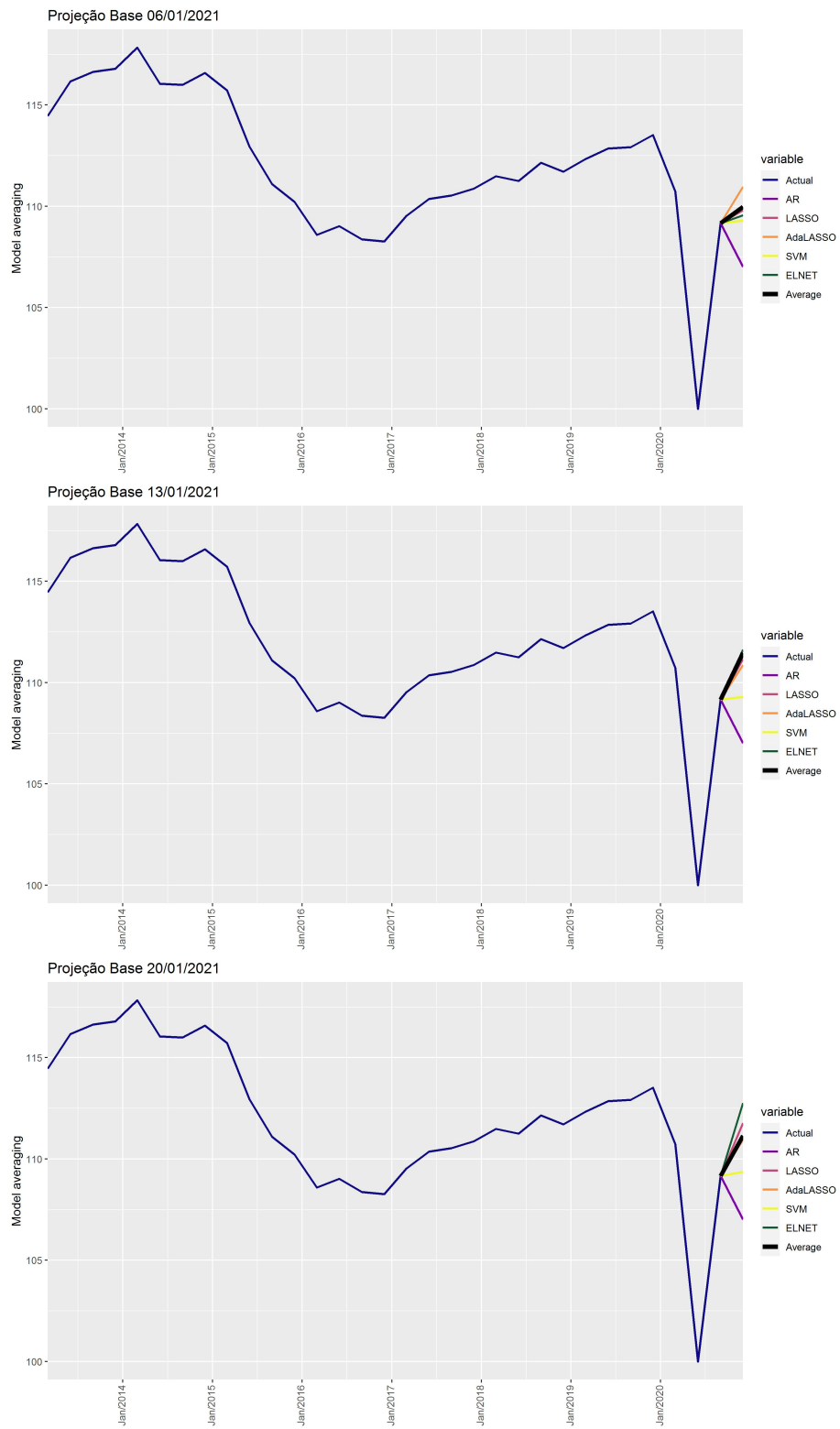


Figura 5.2: Nowcast para algumas datas (1)

5.3 | EXEMPLO DE NOWCAST

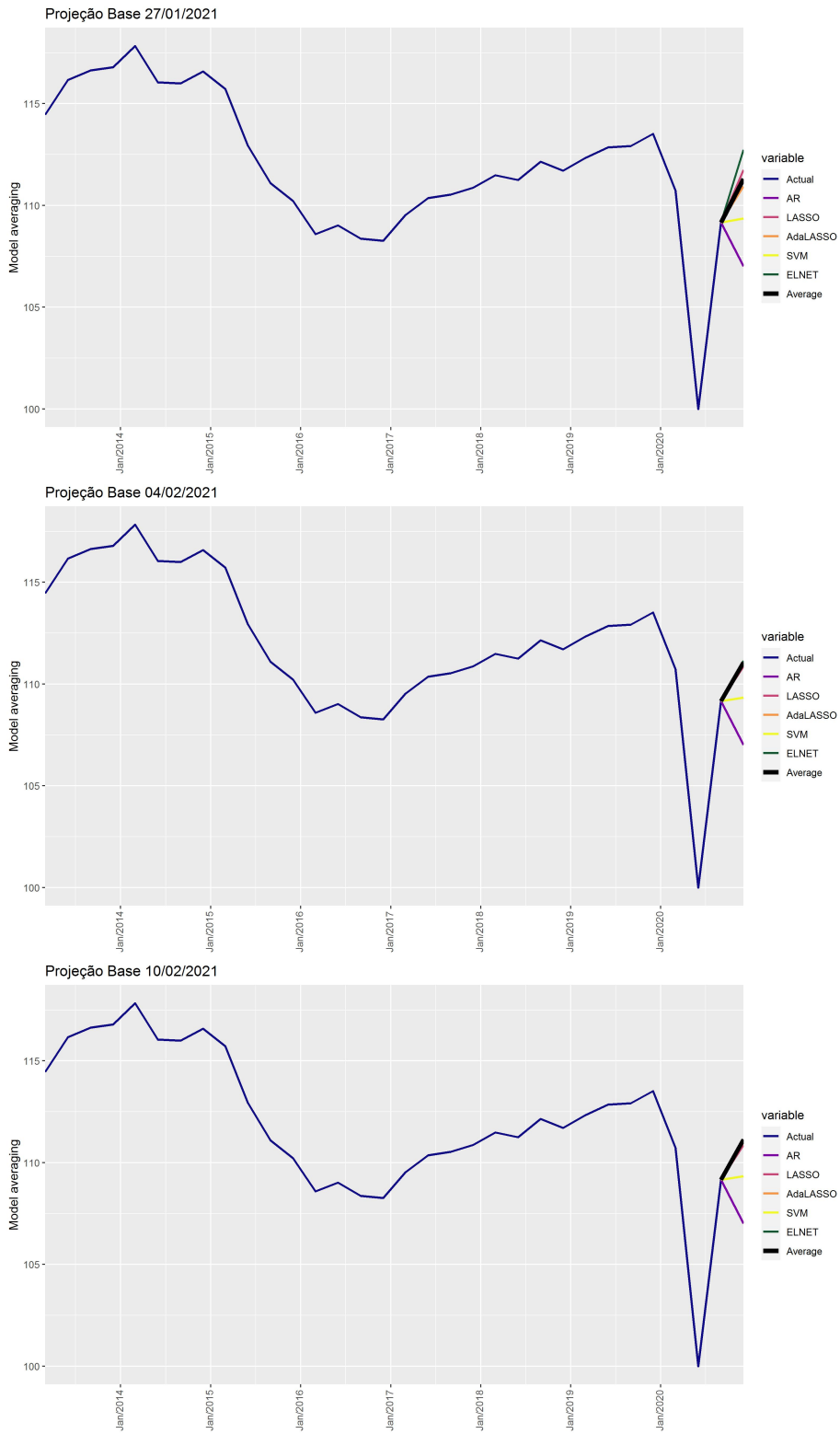


Figura 5.3: Nowcast para algumas datas (2)

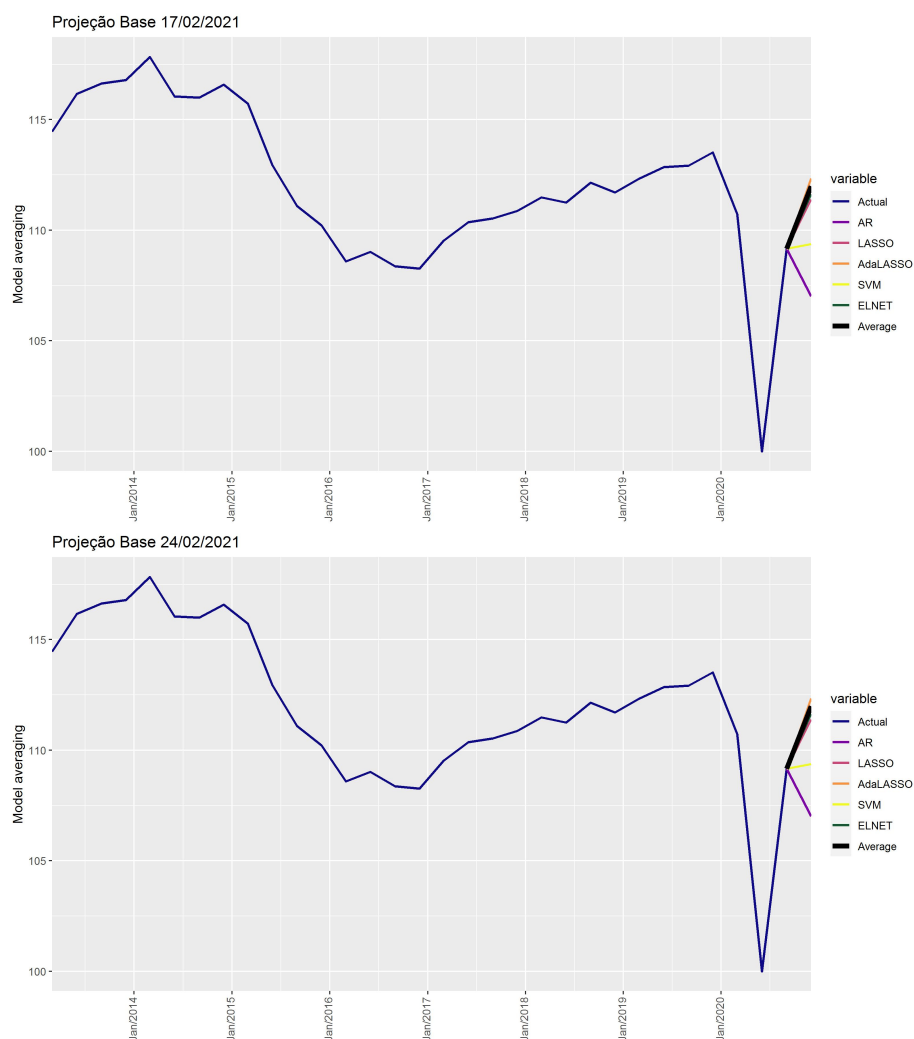


Figura 5.4: Nowcast para algumas datas (3)

## 5.4 Sugestões para pesquisas

Como possíveis extensões deste trabalho podem ser considerados outros métodos de *nowcasting* não abordados (como os métodos FAVARs). Outra possibilidade seria apresentar métodos desagregados, olhando para cada componente do PIB separadamente, prevendo-os separadamente, para em seguida montar uma previsão agregada.

Além disso, podem ser feitas previsões e testes similares para outros países com ambiente econômico similar ao brasileiro, como alguns países da América Latina, para comparar com o caso do Brasil. Algumas categorias de controle de quebra e outras formas de escolha de janela temporal também podem ser implementados e comparados.

# Capítulo 6

## Conclusões

Neste trabalho foram feitos diversos exercícios com modelos de *nowcast* para o PIB brasileiro a partir de técnicas de aprendizagem estatística após o tratamento da base de dados por componentes principais. Os melhores modelos foram filtrados por MCS e, em seguida, o modelo final foi feito baseado na média simples dos modelos com um poder de predição estatisticamente igual. A base de dados foi composta com 280 variáveis explicativas mensais e 4 trimestrais (além do próprio PIB) de 2010 até o final de 2020. Foram estimados modelos de *nowcast* em dezenove datas base distintas para que fosse possível comparar o poder preditivo com a adição de novas variáveis explicativas na base. Como modelo para comparação foi escolhido um AR(4) sem nenhuma variável explicativa.

De início podemos observar que todos os modelos performaram melhor ou igual ao modelo base, indicando que técnicas de *nowcast* apresentam um ganho estatisticamente significativo em relação a modelos comuns de séries de tempo. Além disso, os resultados dos erros dos diferentes períodos mostram a relevância das novas atualizações de informações nos modelos, com uma diminuição drástica na medida ao chegar mais perto da data crítica de divulgação do dado. Este resultado está de acordo com estudos anteriores, Castle, Hendry e Kitov mostraram em 2013 que isto ocorria para países desenvolvidos.

Com isso podemos concluir que a adição das atualizações de variáveis com maior frequência ao longo do tempo auxiliam muito na diminuição do erro de previsão de modelos para o PIB, inclusive em um momento de choque global como o que ocorreu durante os anos de 2020-2021. Os resultados neste trabalho são mais uma evidência da importância da utilização de modelos que utilizam variáveis de alta frequência para que seja possível efetuar decisões de alocação de recursos e de políticas públicas de forma mais assertiva antes da divulgação final de um dado por um órgão público.



## Referências

- [AASTVEIT e TROVIK 2012] Knut Are AASTVEIT e Tørres TROVIK. “Nowcasting norwegian gdp: the role of asset prices in a small open economy”. Em: *Empirical Economics* 42 (1 2012). ISSN: 03777332. DOI: [10.1007/s00181-010-0429-9](https://doi.org/10.1007/s00181-010-0429-9).
- [BAÑBURA, GIANNONE, MODUGNO *et al.* 2013] Marta BAÑBURA, Domenico GIANNONE, Michele MODUGNO e Lucrezia REICHLIN. “Now-casting and the real-time data flow”. Em: *ECB Working Paper Series* 1564 (2013).
- [BAÑBURA, GIANNONE e REICHLIN 2010] Marta BAÑBURA, Domenico GIANNONE e Lucrezia REICHLIN. “Large bayesian vector auto regression”. Em: *Journal of Applied Econometrics* 25 (2010).
- [BRAGOLI *et al.* 2015] Daniela BRAGOLI, Luca METELLI e Michele MODUGNO. “The importance of updating: evidence from a brazilian nowcasting model”. Em: *OECD Journal: Journal of Business Cycle Measurement and Analysis* 2015 (1 2015). ISSN: 19952899. DOI: [10.1787/jbcma-2015-5jrtfl958gmp](https://doi.org/10.1787/jbcma-2015-5jrtfl958gmp).
- [CASTLE *et al.* 2013] Jennifer CASTLE, David HENDRY e Oleg KITOV. “Forecasting and nowcasting macroeconomic variables: a methodological overview”. Em: *Economics Series Working Papers* (674 2013).
- [D’AGOSTINO *et al.* 2016] Antonello D’AGOSTINO, Domenico GIANNONE, Michele LENZA e Michele MODUGNO. “Nowcasting business cycles: a bayesian approach to dynamic heterogeneous factor models”. Em: *Advances in Econometrics* 35 (2016). ISSN: 07319053. DOI: [10.1108/S0731-905320150000035014](https://doi.org/10.1108/S0731-905320150000035014).
- [ERICSSON 2017] Neil R. ERICSSON. “How biased are u.s. government forecasts of the federal debt?” Em: *International Journal of Forecasting* 33 (2 2017). ISSN: 01692070. DOI: [10.1016/j.ijforecast.2016.09.001](https://doi.org/10.1016/j.ijforecast.2016.09.001).

- [FERRARA, GUÉGAN *et al.* 2010] Laurent FERRARA, Dominique GUÉGAN e Patrick RAKOTOMAROLAHY. “Gdp nowcasting with ragged-edge data: a semi-parametric modeling”. Em: *Journal of Forecasting* 29 (1-2 2010). ISSN: 02776693. DOI: [10.1002/for.1159](https://doi.org/10.1002/for.1159).
- [FERRARA e MARSILLI 2019] Laurent FERRARA e Clément MARSILLI. “Nowcasting global economic growth: a factor-augmented mixed-frequency approach”. Em: *World Economy* 42 (3 2019). ISSN: 14679701. DOI: [10.1111/twec.12708](https://doi.org/10.1111/twec.12708).
- [GARCIA *et al.* 2017] Márcio G.P. GARCIA, Marcelo C. MEDEIROS e Gabriel F.R. VASCONCELOS. “Real-time inflation forecasting with high-dimensional models: the case of brazil”. Em: *International Journal of Forecasting* 33 (3 2017). ISSN: 01692070. DOI: [10.1016/j.ijforecast.2017.02.002](https://doi.org/10.1016/j.ijforecast.2017.02.002).
- [GIANNONE *et al.* 2005] Domenico GIANNONE, Lucrezia REICHLIN e David SMALL. “Nowcasting gdp and inflation: the real-time informational content of macroeconomic data releases”. Em: *Finance and Economics Discussion Series* 42 (2005).
- [HANSEN *et al.* 2011] Peter R. HANSEN, Asger LUNDE e James M. NASON. “The model confidence set”. Em: *Econometrica* 79 (2 2011). ISSN: 0012-9682. DOI: [10.3982/ecta5771](https://doi.org/10.3982/ecta5771).
- [KAHOGARA 2015] Douglas KAHOGARA. “Avaliando técnicas de nowcasting: uma aplicação do pib brasileiro”. Em: *Dissertação de mestrado - FGV* (2015).
- [LIU *et al.* 2012] Philip LIU, Troy MATHESON e Rafael ROMEU. “Real-time forecasts of economic activity for latin american economies”. Em: *Economic Modelling* 29 (4 2012). ISSN: 02649993. DOI: [10.1016/j.econmod.2012.03.020](https://doi.org/10.1016/j.econmod.2012.03.020).
- [MODUGNO *et al.* 2016] Michele MODUGNO, Bariş SOYBILGEN e Ege YAZGAN. “Nowcasting turkish gdp and news decomposition”. Em: *International Journal of Forecasting* 32 (4 2016). ISSN: 01692070. DOI: [10.1016/j.ijforecast.2016.07.001](https://doi.org/10.1016/j.ijforecast.2016.07.001).
- [PIAZENTINI e VALLS 2015] Heleno PIAZENTINI e Pedro VALLS. “Um estudo sobre os ciclos de negócios brasileiro (1900-2012)”. Em: *FGV EESP - Textos para Discussão* (2015).
- [YIU e CHOW 2011] Matthew S. YIU e Kenneth K. CHOW. “Nowcasting chinese gdp: information content of economic and financial data”. Em: *China Economic Journal* 3 (3 2011). ISSN: 17538963. DOI: [10.1080/17538963.2010.562028](https://doi.org/10.1080/17538963.2010.562028).