UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

COMPARAÇÃO DAS PENALIZAÇÕES DO TIPO LASSO PARA PREVISÃO DAS TAXAS DE CRESCIMENTO DOS ÍNDICES DE INFLAÇÃO

RAIANE PADILHA SILVEIRA

Porto Alegre
2016

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

COMPARAÇÃO DAS PENALIZAÇÕES DO TIPO LASSO PARA PREVISÃO DAS TAXAS DE CRESCIMENTO DOS ÍNDICES DE INFLAÇÃO

RAIANE PADILHA SILVEIRA

Porto Alegre
2016

RAIANE PADILHA SILVEIRA

COMPARAÇÃO DAS PENALIZAÇÕES DO TIPO LASSO PARA PREVISÃO DAS TAXAS DE CRESCIMENTO DOS ÍNDICES DE INFLAÇÃO

Trabalho de conclusão de curso submetido como requisito parcial para a obtenção do grau de Bacharel em Estatística.

Orientador Metodológico
Profº Dr. Hudson Torrent

Porto Alegre
2016

CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Silveira, Raiane
Comparação das penalizações do tipo Lasso para previsão das taxas de crescimento dos índices de inflação IPCA e IGP-M / Raiane Padilha Silveira, - Porto Alegre: Departamento de Estatística/ UFRGS, 2016
Nº de páginas: 34.

Orientador: Hudson Torrent.

Trabalho de conclusão de Curso – UFRGS, Instituto de Matemática e Estatística, Bacharelado em Estatística, Porto Alegre, 2016.
Orientador: Profº Dr. Hudson Torrent

1. Lasso 2. Adalasso 3. WLadalasso. I. Torrent, Hudson. II. Comparação das penalizações do tipo Lasso para previsão das taxas de crescimento dos índices de inflação

Instituto de Matemática e Estatística
Departamento de Estatística

COMPARAÇÃO DAS PENALIZAÇÕES DO TIPO LASSO PARA PREVISÃO DAS TAXAS DE CRESCIMENTO DOS ÍNDICES DE INFLAÇÃO

Raiane Padilha Silveira

Banca examinadora:

Profº Dr. Hudson Torrent
UFRGS

Profº Dr. Flávio Ziegelmann
UFRGS

Bacharel em Estatística Roberto Koch
Vonpar S/A

AGRADECIMENTOS

A trajetória da graduação foi um enorme desafio e esta realização só se faz possível pelas pessoas maravilhosas que cruzaram o meu caminho.

Primeiramente agradeço a Deus pelas oportunidades, por todas às vezes que eu não sabia o que de fato queria fazer e as coisas simplesmente se ajeitaram da melhor forma possível. Agradeço imensamente aos meus pais, Sandra e Carlos, pelo apoio incondicional na vida inteira, pelos créditos emocionais e financeiros que depositaram em mim, nada disso seria possível sem vocês dois, vocês são os melhores presentes que eu poderia receber de Deus, e tudo que eu quero é dar motivos para que vocês também se sintam gratos por me terem nas suas vidas. Obrigatoriamente tenho que agradecer a minha irmã Raíssa, por nunca me negar nada nessa vida, tens muito ainda para aprender na vida, mas já sabe muito de cumplicidade e lealdade, tenho orgulho de ti. Ao meu namorado Rafael, por me acompanhar desde o primeiro semestre, por ser a pessoa que viveu mais de perto esse turbilhão louco de emoções que é a primeira graduação, obrigada por incentivar, entender e participar. Aos meus amigos da vida inteira (Fernanda, Letícia, Andrezza, Iana, Taiani, Guilherme e Alisson), que sempre acreditaram tanto em mim, e aos novos amigos também (Renata e Roger), que entraram no mesmo ritmo e sempre me fizeram ver que eu podia ir mais um pouco a frente. As minhas colegas e amigas (Camila e Angélica), que dividiram essa trajetória comigo, que me salvaram por tantas vezes. As pessoas que cruzaram o meu caminho nos estágios e depois na Vonpar, por cada um acrescentar na minha vida de alguma forma. Por fim, e não menos importante, aos meus professores, especialmente as profª Lisi e a Vanessa, por serem inspiração na carreira acadêmica, ao profº Alvaro, por ser como um porto seguro no curso, se eu estiver no direito de dar um conselho aos estudantes de estatística na UFRGS seria para falarem com o Alvaro, em todas as vezes que me senti perdida no curso, ele me auxiliou, e o maior dos auxílios foi ter me indicado o profº Hudson. Ao meu querido orientador meu agradecimento eterno, por ser atencioso, disponível e ainda agir como motivador.

Realmente tenho muita sorte.

Obrigada por tudo.

Dedico este trabalho aos meus pais, por tudo.

"O sucesso é ir de fracasso em fracasso sem perder o entusiasmo." (Winston Churchill)

RESUMO

No presente trabalho serão utilizados modelos de alta dimensão Lasso, adaLasso, WLadalasso, para redução de dimensionalidade e previsão de índices usuais da inflação brasileira: IPCA e IGP-M. Os modelos são comparados ao método de previsão geralmente utilizado, autoregressivo linear. Os resultados mostram que os modelos na forma Lasso tem erros quadráticos médios e erros médio absolutos menores para previsões de poucos passos a frente, e para horizontes mais longos os modelos autoregressivos produzem erros menores, embora não sejam diferenças significativas. Todos os modelos foram estimados para os dois índices citados anteriormente, e os resultados mostram que as variáveis que determinam as previsões, são diferentes para cada um dos índices, para o IPCA as variáveis determinantes dizem respeito a petróleo e dívida pública e para o IGP-M além das variáveis de dívida pública também compõe a previsão variáveis referentes à moeda nacional. O estudo é uma abordagem da previsão da taxa de crescimento dos índices em relação às variáveis que cobrem o cenário econômico do país.

Palavras-chave: Lasso, AdaLasso, WLadalasso, previsão, séries temporais.

ABSTRACT

In the present work will be used high-dimensional models Lasso, adaLasso, WLadalasso, for dimensionality reduction and prediction of the usual Brazilian inflation indexes: IPCA e IGP-M. The models are compared to the commonly used, autoregressive linear prediction method. The results show that the models in the Lasso form have mean square errors and mean absolute errors smaller for predictions of a few steps forward, and for longer horizons the autoregressive models produce smaller errors, although they are not significant differences. All models were estimated for the two indexes mentioned above, and the results show that the variables that determine the predictions are different for each index, for the IPCA the determinant variables concern oil and public debt and for the IGP-M besides the public debt variables also composes the forecast variables referring to the national currency.
The study is an approach to predict the index growth rate in relation to the variables that cover the country's economic scenario.

Keywords: Lasso, AdaLasso, WLadalasso, Forecast, time series.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

Introdução

Os índices de inflação são utilizados para medir a variação dos preços e o impacto no custo de vida da população. Existem diversos índices importantes de inflação, em que a medição é feita por órgãos especializados do governo. Iremos trabalhar com IPCA e IGP-M. O IPCA que é o índice de preços oficial brasileiro, é calculado mensalmente pelo IBGE usando dados de famílias que ganham entre 1 e 40 salários mínimos e vivem em áreas urbanas. E o IGP-M, medido pela FGV, registra inflação dos preços desde matérias primas agrícolas industriais até bens e serviços finais.

Além de ser uma variável importante para todos os âmbitos sociais, a inflação é uma variável difícil de prever, principalmente no caso do estudo, em que estaremos prevendo a inflação brasileira, uma vez que o país vive constantes momentos de turbulência econômica. Ao construirmos uma previsão confiável desta variável estaremos dando maior segurança às futuras decisões econômicas.

Os modelos usuais de Box e Jenkins (1970) nos fornecem uma previsão com base somente nos dados passados dos índices estudados, o que nos leva a contestar o desprezo de outras variáveis que podem vir a ser importantes para a análise.

Ao pensarmos em quais variáveis podem compor esta previsão nos deparamos com um leque muito grande de informação e o grande problema é determinar que variáveis são essas, considerando acurácia, complexidade computacional, e sem sacrificar as interpretações, principalmente quando temos uma amostra pequena. Talvez por este motivo nos deparamos com vários indicadores diferentes, e não uma unificação dos preditores importantes em um único indicador. Uma primeira alternativa para solução da inclusão de possíveis variáveis que possam vir a ser importantes para previsão é a analise fatorial, detalhada por Medeiros et al. (2016), entretanto não replicaremos esta parte do trabalho, por temos perda na interpretação, uma vez que não sabemos quais variáveis são de fato importantes para o modelo.

Diante das necessidades encontradas e do crescimento na literatura dos modelos de alta dimensionalidade, analisando o ganho substancial desta utilização para previsão, incorporaremos o mesmo neste estudo. Tibshirani (1996) propôs o Lasso, em um mesmo contexto de regressão linear, impondo uma penalização aos coeficientes. Essa penalização força que alguns coeficientes se reduzam a zero, reduzindo a dimensionalidade e o espaço de parâmetros. O Lasso é atualmente uma usual técnica de regularização, funcionando essencialmente nos casos em que existem muitos coeficientes nulos dentro do conjunto de coeficientes a serem estimados, com o objetivo de selecionar as variáveis preditoras que melhor explicam a variável resposta e construir previsões com menor variância.

Zou (2006) estuda as propriedades do Lasso, mencionadas por Fan e Li (2001) e mostra situações em que o Lasso não é consistente, propondo então o Lasso Adaptativo (AdaLasso), penalizando os coeficientes com diferentes pesos, solucionando os problemas de consistência do Lasso, também seguindo o estudo de Medeiros et al. (2016) utilizaremos os pesos 0,3 para IPCA e 0,5 para IGP-M.

Konzen e Ziegelmann (2016) tem o objetivo de investigar como a penalidade sobre o conjunto de coeficientes pode contribuir no desempenho das previsões, ao apresentar o WLadalasso, que além de considerar diferentes pesos para os coeficientes penaliza também pela defasagem, lags mais elevados são mais penalizadas, uma vez que uma informação mais recente pode ser mais importante para previsão. Através de simulações e uma aplicação financeira os autores mostram que por uma dependência linear forte entre preditores e um maior número de defasagens o WLadalasso supera os outros métodos de penalização na perspectiva de seleção das covariáveis e estimativas de previsão em horizontes menores.

Este trabalho é uma releitura de Medeiros et al. (2016), e a extensão segue a mesma técnica de Konzen e Ziegelmann (2016), com os índices IPCA e IGP-M. Tem o objetivo de prever a taxa de crescimento da inflação e comparar as técnicas descritas anteriormente em diferentes horizontes de previsão, além de observar as variáveis selecionadas como preditoras. Além disso, em termos macroeconômicos, desejamos investigar a relação dos modelos escolhidos frente à Curva de Phillips, que é uma das relações mais usuais para estruturar previsões de inflação.

A seguir faremos a revisão das definições dos modelos que serão utilizados: ARIMA, Lasso, AdaLasso e WLadalasso. Seguiremos com a definição formal da variável a ser prevista e nos seguintes capítulos estarão as análises, resultados e comparações obtidas.

1. Lasso

O Lasso (least absolute shrinkage and selection operador), Tibshirani(1996), é um método de encolhimento do conjunto de coeficientes. Tem como objetivo estimar um modelo que determine o conjunto de preditores que melhor expliquem a variável resposta, e produza previsões com pequena variância.

As estimativas são obtidas através da minimização dos quadrados dos resíduos, sujeita à penalização da norma L1 (segundo termo em (1)) dos coeficientes, reduzindo os coeficientes irrelevantes a zero.

$$\hat{\beta} = \arg min_{\hat{\beta}} \; ||Y - X\beta||_2^2 + \lambda \sum_{j=1}^{p} |\beta_j| \qquad (1)$$

β é um vetor nx1, Y = $(y_1, \dots, y_n)$ é a variável resposta, X é uma matriz pxn com as variáveis preditoras e λ é o parâmetro de encolhimento, quando λ = 0 as estimativas são as mesmas que MQO. λ é estabelecido por validação cruzada, esta técnica particiona a amostra original em K subamostras de tamanho igual (neste trabalho K=10), uma das K subamostras é retirada e o modelo é estimado com K-1 subamostras.

A figura 1 mostra o efeito de λ sobre as estimativas dos coeficientes, cada linha representa o valor do coeficiente de uma variável diferente. O número de variáveis selecionadas e o tamanho dos coeficientes diminuem à medida que aumenta o λ.

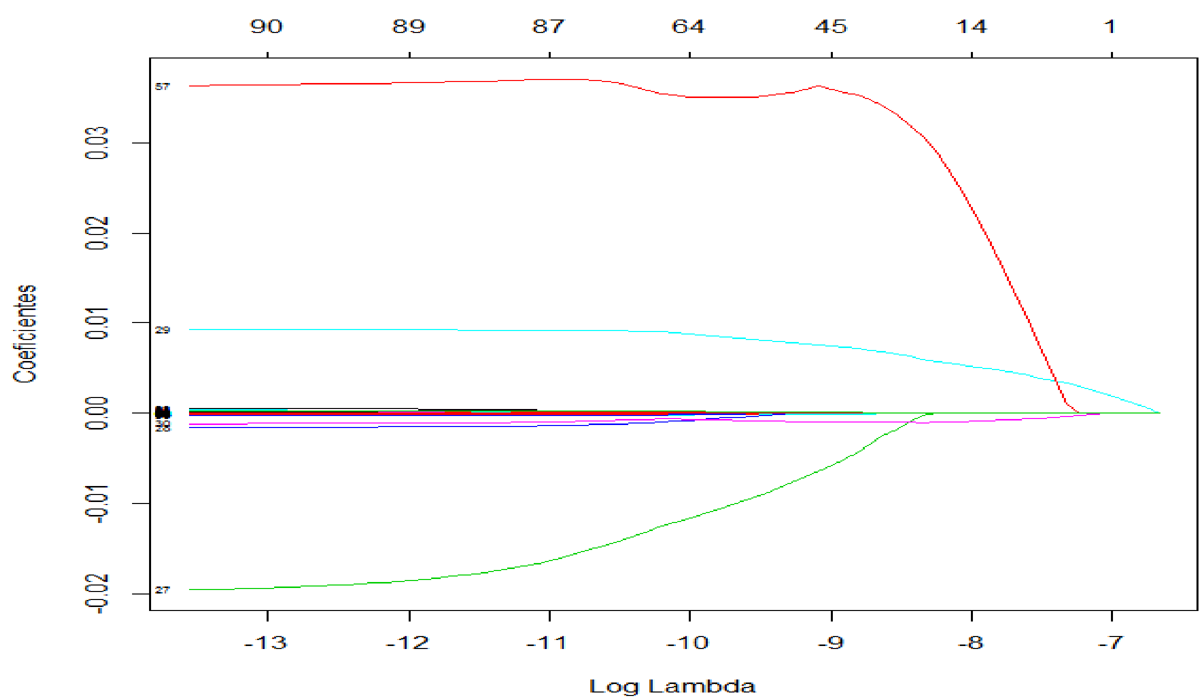

Figura 1- λ vs Coeficientes

(Elaborado por Medeiros et al. (2016))

Considerando consistência dos modelos Lasso, é importante saber se a solução representa bem o modelo isso nada mais é do que gozar das propriedades Oracle, ou seja, conseguir selecionar o subconjunto correto de variáveis relevantes e seus parâmetros devem ser assintoticamente como dos estimadores MQO, o que só acontecerá se o subconjunto de variáveis escolhido for o das variáveis relevantes para o modelo. De acordo com Konzen (2014, p. 21)

> Zhao e Yu (2006) consideram dois problemas: 1) se existe uma quantidade determinística de regularização que fornece consistência de seleção; 2) se para cada amostra existe uma quantidade correta de regularização que seleciona o melhor modelo. Estes resultados mostram que existe uma condição que denominam por "Condição Irrepresentável" que é quase necessária e suficiente para ambos os tipos de consistência. Estes resultados são válidos para modelos lineares.

Neste trabalho, o principal ganho na utilização do Lasso, é a possibilidade de analisar inflação com base em outras variáveis que podem ser consideradas importantes para sua previsão. Quando consideramos trabalhar com a classe de modelos ARIMA, descartamos informações de demais variáveis que podem ser importantes para previsão, e a analise fatorial é desconsiderada em função do tamanho da nossa amostra impossibilitar sua execução.

2. AdaLasso

Existem situações em que o Lasso é inconsistente para seleção de variáveis, com base nisso Zou(2006) propôs o AdaLasso. O AdaLasso se difere do Lasso ao aplicar pesos diferentes aos coeficientes. Os pesos servem para auxiliar na seleção de variáveis mais importantes, quanto maior o coeficiente da variável menor é o peso atribuído a mesma, sendo assim ficam mais penalizadas as variáveis menos importantes.

$$\hat{\beta} = \arg min_{\hat{\beta}} \; ||Y - X\beta||_2^2 + \lambda \sum_{j=1}^{p} w_j |\beta_j| \tag{2}$$

Onde $wj = |\widehat{\beta_j^*}|^{-\tau}$, $|\widehat{\beta_j^*}|$ é uma estimativa do primeiro passo (Lasso) e τ é m parâmetro escolhido usando o mesmo critério de λ, para o presente trabalho τ será o mesmo utilizado em Medeiros et al. (2016).

As definições de consistência estão detalhadas em Konzen e Ziegelmann (2016). Em resumo, temos que, o adaLasso seleciona corretamente as variáveis relevantes quando o tamanho da amostra aumenta, e possui estimativa dos coeficientes não-nulos com assintoticamente a mesma distribuição que os estimadores MQO, quando estimado apenas com as variáveis relevantes.

3. WLadalasso

O grande ganho na utilização do WLadalasso (Weighted Lag adaptative Lasso) é considerar cada uma das defasagens das variáveis preditoras. Proposto por Konzen(2014) questiona se a penalização menor para lags menos distantes contribui para previsão, uma vez que uma informação mais recente pode ser mais relevante.

$$\hat{\beta} = \arg min_{\hat{\beta}} \; ||Y - X\beta||_2^2 + \lambda \sum_{j=1}^{p} w_j |\beta_j| \tag{3}$$

Onde $wj = (|\widehat{\beta_j^*}| e^{-\alpha l})^{-\tau}$, $\tau > 0,\ \alpha \geq 0$ *e* l é a ordem da defasagem.

Konzen e Ziegelmann (2016) mostram através de resultados de simulações e de um estudo sobre o PIB que a penalização diferente para cada defasagem contribui para seleção das variáveis mais importantes e para o desempenho da previsão, com resultados superiores comparados aos métodos Ridge, Lasso e adaLasso, principalmente em pequenas amostras. Nas simulações foram testados casos com poucas, algumas e muitas variáveis relevantes e variando a importância do lag.

Nos capítulos a seguir o WLadalasso será comparado com os resultados de Lasso, AdaLasso, AR.

4. ARIMA – Modelos Autoregressivos Integrados e de Médias Móveis

A metodologia Box-Jekins identifica o comportamento da correlação seriada ou autocorrelação dos valores de uma série temporal, isto é, de um conjunto de pontos observados ao longo do tempo.

Estes modelos são univariados e mais tradicionalmente utilizados para previsões de curto prazo, entretanto são ineficientes em séries temporais com comportamentos mais complexos e recebem críticas quanto à subjetividade da escolha do modelo, mesmo que existam testes estatísticos para verificar a validade dos modelos e medidas de incerteza das previsões.

A primeira exigência para modelagens, é que a série temporal seja estacionária, ou possa ser transformada em estacionária. A série é dita estacionária quando a média, variância e estrutura de autocorrelação não mudam no decorrer do tempo. Um processo estocástico que apresenta média e variância constante ao longo do tempo e função de autocovariância que depende somente da defasagem entre os instantes de tempo é definido como fracamente estacionário (ou estacionário em covariância). Um processo estritamente estacionário é àquele que apresenta a distribuição de X_t , independente de t. Para um processo X_t ser considerado estacionário, as seguintes condições devem estar satisfeitas:

I. $E(X_t) = \mu$;

II. $E(X_t^2) < \infty$;

III. $Cov[X(t), X(t+1)] = \gamma(\tau)$

A transformação mais comum para tornar uma série estacionária consiste em tomar diferenças sucessivas da série original, até obter a estacionariedade. A primeira diferença Z(t) é definida por:

$$\Delta Z(t) = Z(t) - Z(t-1) \qquad (4)$$

Neste trabalho serão apresentados os gráficos de ACF e PACF para ilustração de estacionariedade, entretanto vamos considerar os resultados de Medeiros et al. (2016) que utilizou o ADF teste que apresenta a hipótese nula como condição de estacionariedade para justificar a tomada da primeira diferença da série.

Na figura 2 observamos variância constante, e pela figura 3 no gráfico de ACF observamos autocorrelações, para visualmente iniciar a modelagem. Uma ACF com grandes picos nos lags iniciais que declinam para zero ou PACF om grande pico no primeiro lag, possivelmente indica modelagem de um processo autoregressivo. Uma ACF com grandes picos no primeiro e segundo lag e uma PACF com grandes picos nos lags iniciais indicam um processo de média móvel. Quando o gráfico de ACF e PACF tem grandes picos que declinam gradualmente, temos um indicativo de modelagem ARMA.

Os procedimentos desta metodologia são resumidamente descritos em três passos:

1) Identificação: consiste em encontrar o modelo que descreve de forma mais adequada a série temporal, verificando estacionariedade;

2) Estimação: estimar os parâmetros do modelo escolhido.

3) Verificação: estudo da adequabilidade do modelo estimado, sendo feita através dos resíduos.

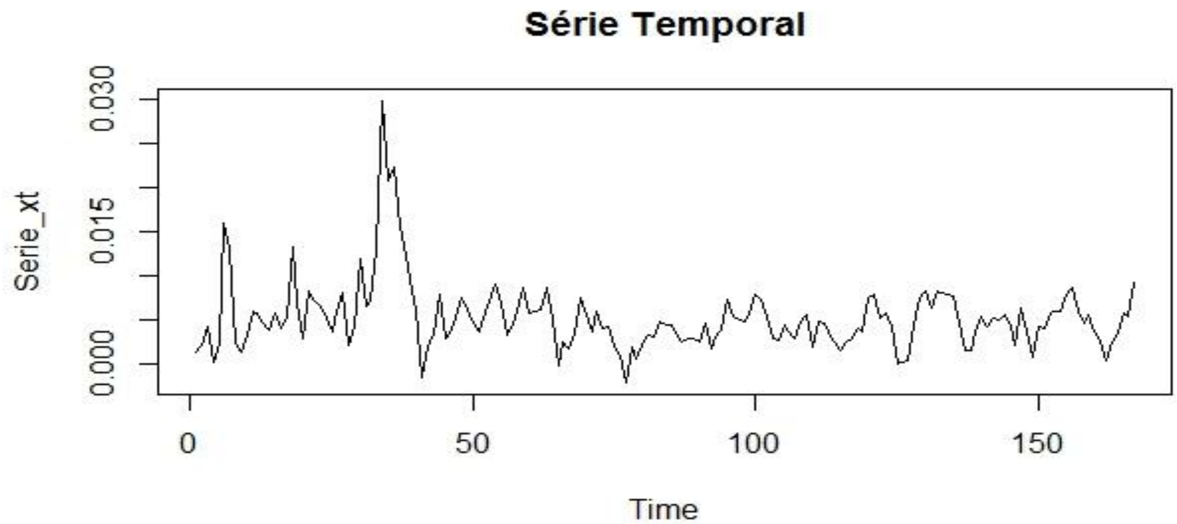

Figura 2- Série Temporal Estacionária

(Elaborado pelo Autor)

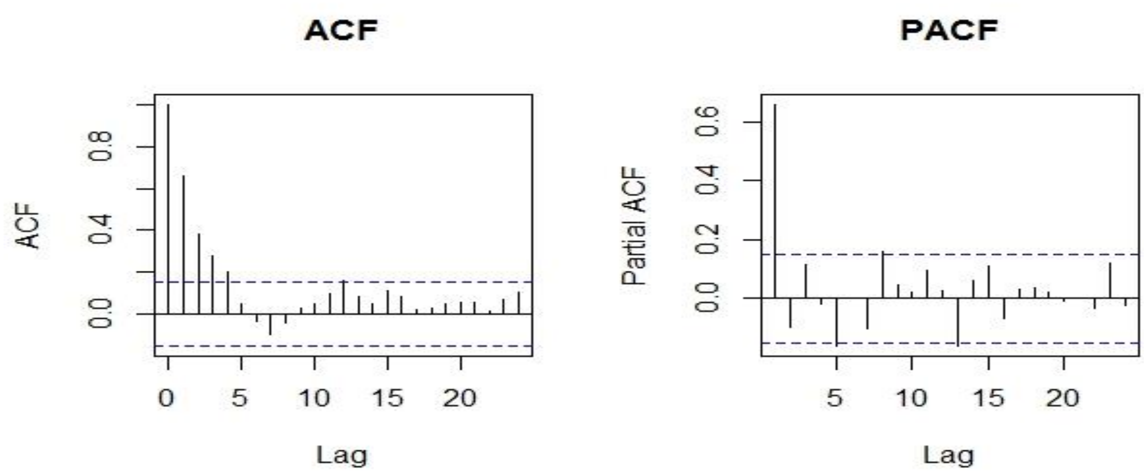

Figura 3 - Análise Estacionariedade

(Elaborado pelo Autor)

Dada X_t uma série estacionária, existem três modelos básicos que podem ser assumidos, como já mencionado: AR (autoregressivo), MA (médias móveis) e um combinados de ambos, ARMA. Quando for aplicada uma diferenciação regular junto com a combinação dos modelos, então a referência é dada por ARIMA. Essa composição é popularmente chamada de “filtro”, podendo a série ser modelada apenas com alguns “filtros” ou com toda totalidade.

Dizemos que X_t é um processo autoregressivo de ordem p, $(X_t \sim AR(p))$, se podemos escrever o processo na seguinte forma:

$$X_t = \Phi_o + \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + \varepsilon_t \quad (5)$$

Se X_t é um processo média móveis de ordem q, $(X_t \sim MA(q))$, se podemos escrever o processo na seguinte forma:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad (6)$$

Quando X_t é um processo autoregressivo e de média móveis de ordem (p,q), $(X_t \sim ARMA(p,q))$, se podemos escrever o processo na seguinte forma:

$$X_t = \Phi_o + \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad (7)$$

5. Curva Phillips

A importância teórica e empírica do estudo da Curva de Phillips é comprovada pela utilização da maioria dos bancos centrais ao redor do mundo. É a correlação negativa entre inflação e desemprego, segundo Phillips (1958), quanto maior o desemprego, menor a inflação e vice-versa.

A utilização da Curva Phillips é um referencial em termos de previsão de inflação, existe um forte debate econômico sobre a relevância e precisão da mesma. Diante disso passamos a questionar nossos resultados que não traziam nenhuma variável relacionada ao desemprego.

Pretendemos encontrar relações ou diferenças dos nossos resultados em relação a este referencial.

6. Aplicações

A taxa de inflação é resumidamente a média do crescimento dos preços de um conjunto de bens e serviços em um determinado período. Sendo medida através dos índices de preço, neste trabalho consideraremos o IPCA (Índice Nacional de Preços ao Consumidor Amplo) e o IGP-M (Índice Geral de Preços do Mercado), o primeiro pondera os bens e serviços mais importantes para população e medem o crescimento desses preços, e o segundo é um indicador para balizar as correções de títulos emitidos pelo Tesouro Nacional e Depósitos Bancários com renda pós-fixada acima de um ano. Existem outros indicadores de inflação e os mesmos serão utilizados como variáveis preditoras nos capítulos a seguir.

Nas subseções a seguir vamos comparar as previsões das variáveis em estudo em diferentes horizontes e pelos modelos citados anteriormente. O objetivo desta seção é encontrar os âmbitos de ganho de cada uma das aplicações, a fim de construir previsões com menores erros dos índices de inflação. Todos os resultados foram construídos no software R Project, com auxílio do pacote glmnet.

O conjunto de dados é o mesmo de Medeiros et al. (2016), as 95 variáveis preditoras foram coletadas do Banco Central do Brasil, FGV, IBGE, IPEADATA, e o banco de dados Bloomberg, no período de janeiro de 2000 a dezembro de 2013 e cobrem índices de produção, dívida pública, índices de preços, mercados financeiros, impostos, importação e exportação, contas do governo, investimentos, salários e variáveis internacionais que podem estar relacionadas a economia brasileira. A escolha deste período é porque a política de metas de inflação começou no Brasil em julho de 1999, além da adaptação a nova moeda, implementada em 1994.

Para os dois índices foram estimados os modelos: AR(1), Lasso, adaLasso tendo o Lasso como primeira etapa, e WLadalasso como extensão do adaLasso, acrescendo das defasagens. Além das comparações dos erros de predição, as mesmas ainda serão comparadas pelo teste de Giacomini e White (2006), este teste compara se duas previsões são estatisticamente diferentes.

As previsões foram estimadas usando janela de rolamento de 132 observações para t={2,...,12} horizontes, cada passo previsto foi utilizado para próxima previsão. O período fora da amostra vai de janeiro de 2011 a dezembro de 2013. As previsões foram comparadas com o teste de Giacomini e White (2006).

a. IPCA

A primeira aplicação na modelagem dos dados foi univariada, com modelos da classe ARIMA. Na figura 4 observamos variância constante, e pela figura 5 no gráfico de ACF observamos autocorrelações, visualmente temos AR(1).

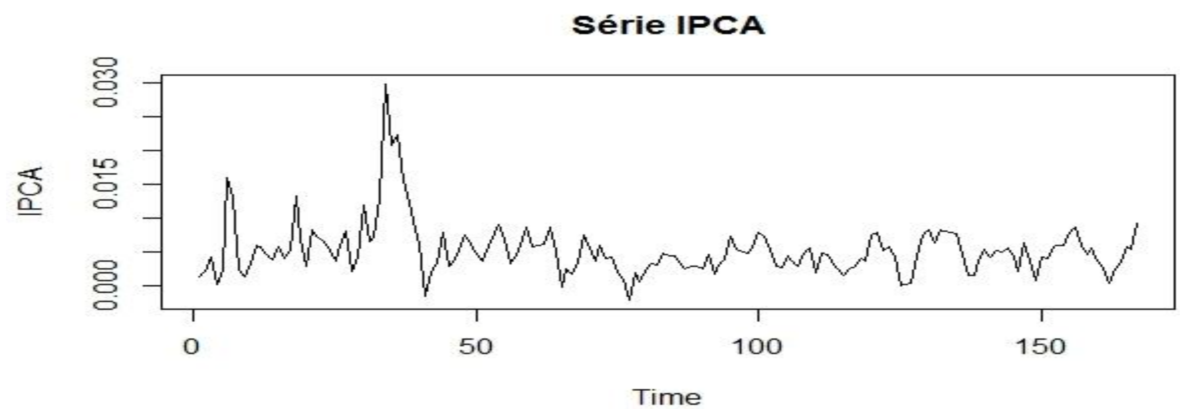

Figura 4 – IPCA

(Elaborado pelo Autor)

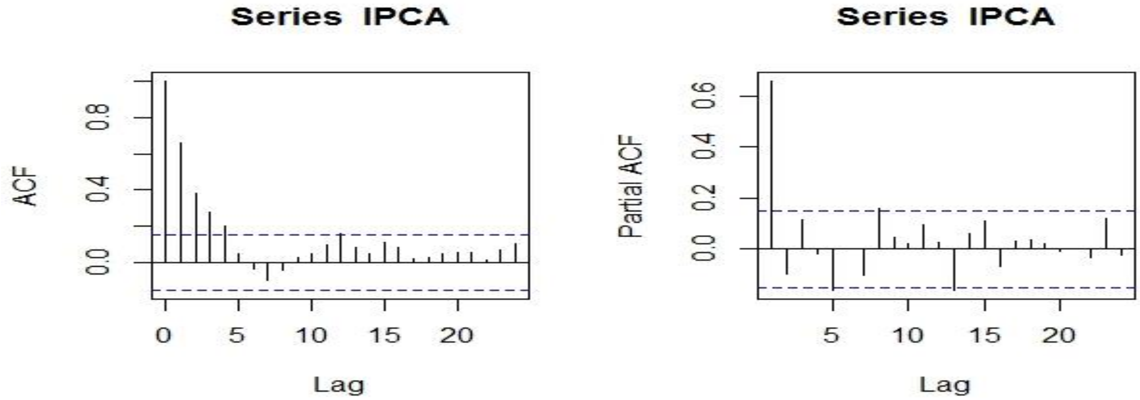

Figura 5 - Análise ACF / PACF

(Elaborado pelo Autor)

Pelos critérios de AIC, significância dos parâmetros e pelos resultados do Box-Test – O Box Test tem como hipótese nula a condição dos resíduos serem ruído branco, condição fundamental da análise dos modelos da classe ARIMA – todos os horizontes de previsão foram modelados com AR(1).

Sabe-se que apesar dos modelos AR(1) considerar apenas o passado da variável para construção da previsão os resultados normalmente são bons para séries bem comportadas. Para concluirmos a respeito disso, mais a frente consideraremos raiz do erro quadrático médio e erro médio absoluto. A grande perda é a falta de interpretação que os demais coeficientes podem trazer a respeito da previsão.

A segunda modelagem é feita através do Lasso, para exemplificar a construção do modelo e da previsão vamos considerar a modelagem 2 passos a frente. Todas as variáveis explicativas foram diferenciadas para atingir estacionariedade, pois quando realizado o teste ADF a hipótese nula de raiz unitária foi rejeitada para todas as variáveis.

Pela figura 6, vemos os dois log(λ) sugeridos com base nos dois critérios mais habituais: o primeiro, a esquerda, é o log(λ) que proporciona o erro médio mínimo (-7.3753), e o segundo é o log(λ) que fornece o modelo mais regularizado (-6,6310).

Ao utilizarmos o λ que fornece o erro médio mínimo, temos um modelo com intercepto e mais 5 variáveis explicativas, são referentes a consumo de petróleo, consumo de derivados do petróleo, dividas interna dos municípios, estados e governo e taxa de juros do banco central. Sendo assim, com base nessa primeira análise concluiríamos que o que afeta imprescindivelmente a taxa de crescimento do IPCA são os indicadores de petróleo e as dividas internas do país.

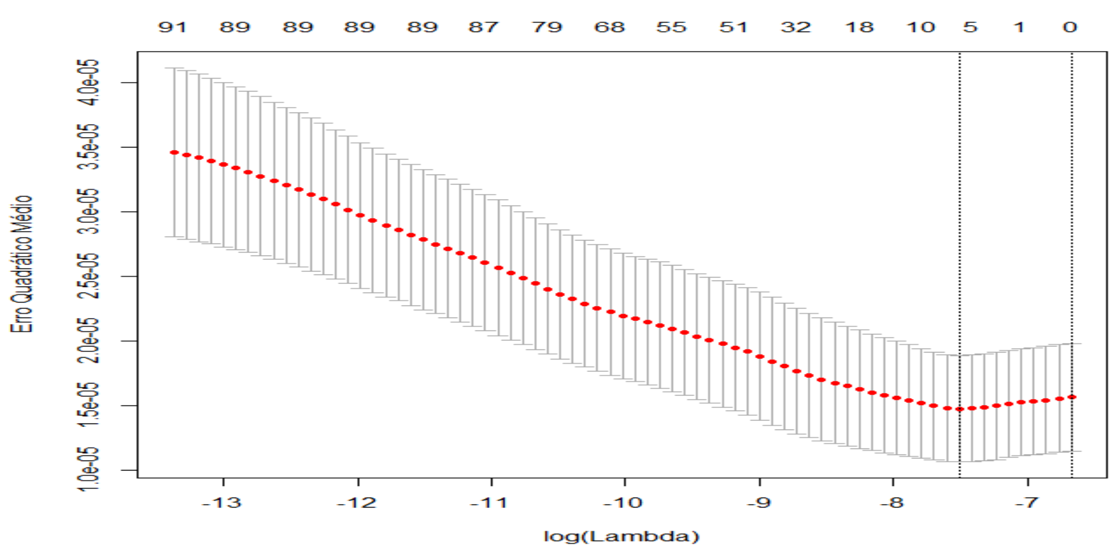

Figura 6 – Escolha λ

(Pacote glmnet – R Project)

A terceira modelagem é feita através do adaLasso, como no Lasso todas as variáveis explicativas foram diferenciadas, o λ foi utilizado o mesmo do procedimento anterior e a escolha de τ foi por validação cruzada.

Temos um modelo com as mesmas variáveis explicativas do anterior, mas com menor erro, como veremos a seguir. Em comparação com o Lasso o ganho é quanto aos erros de previsão.

A quarta modelagem, feita através do WLadalasso, tendo o adaLasso acrescido das defasagens, a ordem da defasagem de cada horizonte foi escolhido pelo critério AIC, foram consideradas opções de lag ℓ = (5, 10, 15, 20) seguindo as mesmas escolhas de Konzen e Ziegelmann (2016). O modelo é reduzido para 4 variáveis explicativas, retirando do modelo a variável de taxa de câmbio, neste modelo dois passos a frente o WLadalasso nos fornece o menor erro de previsão. Demonstrando que o acréscimo da defasagem é positivo para previsão.

Em geral, em relação à seleção de variáveis, em todos os horizontes o comportamento foi o mesmo. Quanto ao modelo escolhido para cada horizonte, podemos considerar as figuras 7 e 8 que mostram a raiz dos erros quadráticos médios e erros médio absoluto das previsões. De t+2 até t+6 o WLadalasso fornece superioridade preditiva em relação a raiz do erro quadrático médio, em t+7 até t+12 o melhor desempenho quanto a este erro é da modelagem AR(1). O erro médio absoluto tem comportamento semelhante a raiz do erro quadrático médio, temos que em t+2 até t+7 o WLadalasso é superior, e de t+7 até t+12 a modelagem AR(1) apresenta erros menores.

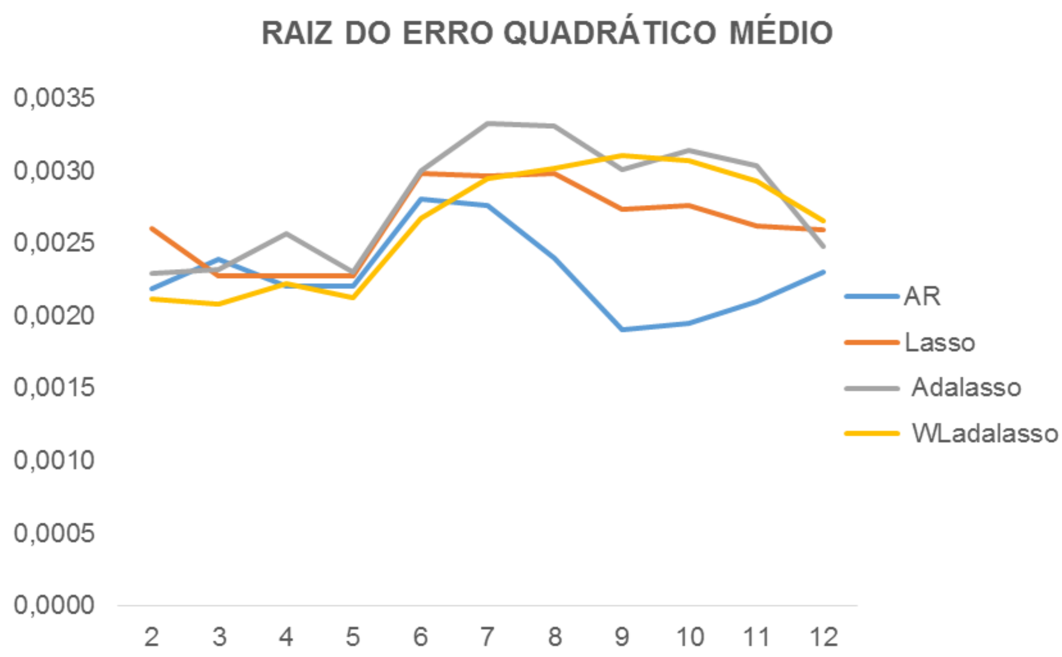

Figura 7- IPCA RAÍZ DO ERRO QUADRÁTICO MÉDIO

(Baseado em Medeiros et al. (2016))

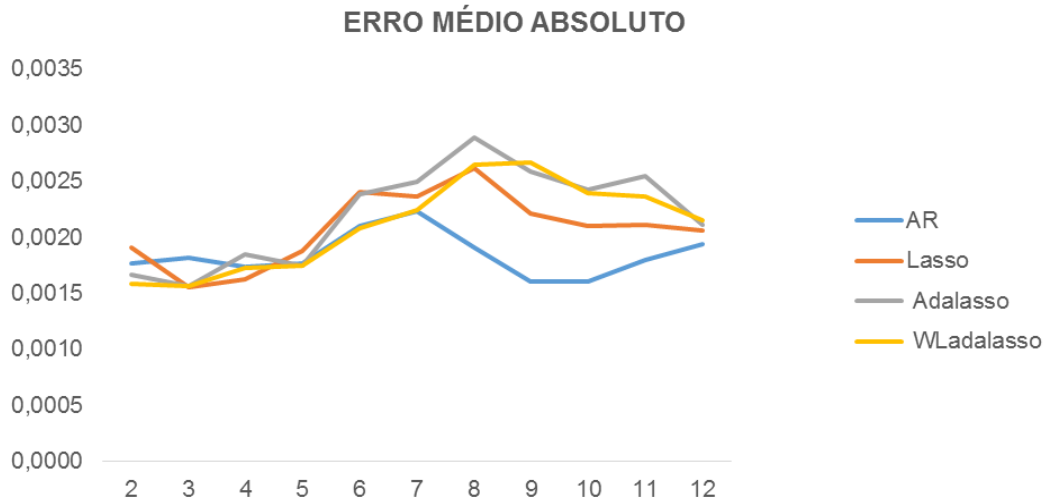

Figura 8 - IPCA – ERRO MÉDIO ABSOLUTO

(Baseado em Medeiros et al. (2016))

A Tabela 1 reflete o GW-test (Giacomini e White), os p-valores verificam se as previsões são significativamente iguais. Os resultados mostram que apesar dos ganhos já apresentados anteriormente, em relação aos erros de predição, o ganho no desempenho não é estatisticamente representativo.

Diante destes resultados, uma vez que as previsões são estatisticamente semelhantes, as previsões do modelo WLadalasso podem ser consideradas pelo menos tão boas, quanto a melhor previsão de cada horizonte, uma vez que a modelagem é mais consistente. Este resultado sustenta a evidencia de que as variáveis selecionadas, são de fato relevantes para gerar os horizontes menores de previsão da taxa de crescimento do IPCA, sendo assim o WLadalasso resolve o problema da alta dimensionalidade do espaço paramétrico e oferece poder preditivo.

Em geral as variáveis selecionadas são: consumo de petróleo, consumo de derivados do petróleo, dividas interna dos municípios, estados e governo e taxa de juros do banco central. Isso mostra que ao modelarmos AR(1), estamos negligenciando o uso destas informações importantes. Por fim, nossos resultados sugerem que as quatro últimas variáveis são incorporadas nas previsões do índice IPCA, ao olharmos todos horizontes.

a) **Consumo de petróleo**: A alta do preço do petróleo impacta no consumo e contribui para o aumento da inflação.
b) **Consumo de derivados do petróleo**: Com a alta do petróleo, o mesmo acontece com os derivados, resultando no mesmo impacto anterior.
c) **Dívidas internas dos municípios, estados e governos**: O crescimento dos endividamentos municipais, estaduais e governamentais acarreta no aumento da inflação.
d) **Taxa de juros**: As taxas de juros interferem no volume de investimento, estando conectada com a inflação.

t+2				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,2562	0,1514	0,2005
Lasso	0,2562	-	0,0927	0,1800
Adalasso	0,1514	0,0927	-	0,3190
Wladalasso	0,2005	0,1800	0,3190	-

t+3				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9492	0,7038	0,9999
Lasso	0,9492	-	0,5117	0,5251
Adalasso	0,7038	0,5117	-	0,5005
Wladalasso	0,9999	0,5251	0,5005	-

t+4				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,3893	0,5952	0,4917
Lasso	0,3893	-	0,9648	0,7469
Adalasso	0,5952	0,9648	-	0,2984
Wladalasso	0,4917	0,7469	0,2984	-

t+5				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,7045	0,6013	0,6063
Lasso	0,7045	-	0,2423	0,1241
Adalasso	0,6013	0,2423	-	0,5028
Wladalasso	0,6063	0,1241	0,5028	-

t+6				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,8854	0,8798	0,7980
Lasso	0,8854	-	0,4690	0,0035
Adalasso	0,8798	0,4690	-	0,1372
Wladalasso	0,7980	0,0035	0,1372	-

t+7				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,8616	0,9208	0,8119
Lasso	0,8616	-	0,3365	0,1556
Adalasso	0,9208	0,3365	-	0,2226
Wladalasso	0,8119	0,1556	0,2226	-

t+8				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9963	0,9731	0,9752
Lasso	0,9963	-	0,8202	0,5630
Adalasso	0,9731	0,8202	-	0,1695
Wladalasso	0,9752	0,5630	0,1695	-

t+9				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,8850	0,9230	0,9998
Lasso	0,8850	-	0,9100	0,9900
Adalasso	0,9230	0,9100	-	0,6102
Wladalasso	0,9998	0,9900	0,6102	-

t+10				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9781	0,9962	0,9982
Lasso	0,9781	-	0,9729	0,9840
Adalasso	0,9962	0,9729	-	0,4081
Wladalasso	0,9982	0,9840	0,4081	-

t+11				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9961	0,9983	0,9999
Lasso	0,9961	-	0,9931	0,9760
Adalasso	0,9983	0,9931	-	0,1769
Wladalasso	0,9999	0,9760	0,1769	-

t+12				
Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,7427	0,7216	0,7822
Lasso	0,7427	-	0,5652	0,7647
Adalasso	0,7216	0,5652	-	0,5524
Wladalasso	0,7822	0,7647	0,5524	-

Tabela 1 – Esta tabela mostra o p-valor do GW-test (a hipótese nula do GW é de que as previsões são estatisticamente iguais, (Baseado em Medeiros et al. (2016))

b. IGP-M

O IGP-M é mais volátil que o IPCA, seu coeficiente de variação é 1,24 e o do IPCA é 0,73. A correlação entre os dois índices é 0,73, o que pode ser resultado dos itens que constituem as cestas de bens de cada índice, uma vez que são diferentes.

Como no IPCA as previsões foram calculadas para t={2,...,12} horizontes. Considerando a modelagem univariada, com modelos da classe ARIMA. Na figura 9 observamos variância constante, e pela figura 10 no gráfico de ACF observamos autocorrelações, nota-se necessidade da utilização do modelo autorregressivo de ordem 1.

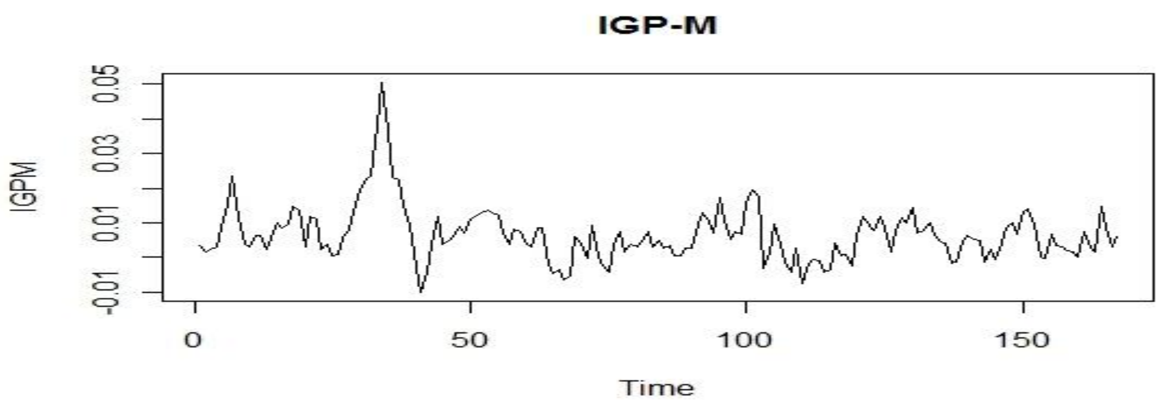

Figura 9 - IGP-M

(Elaborado pelo Autor)

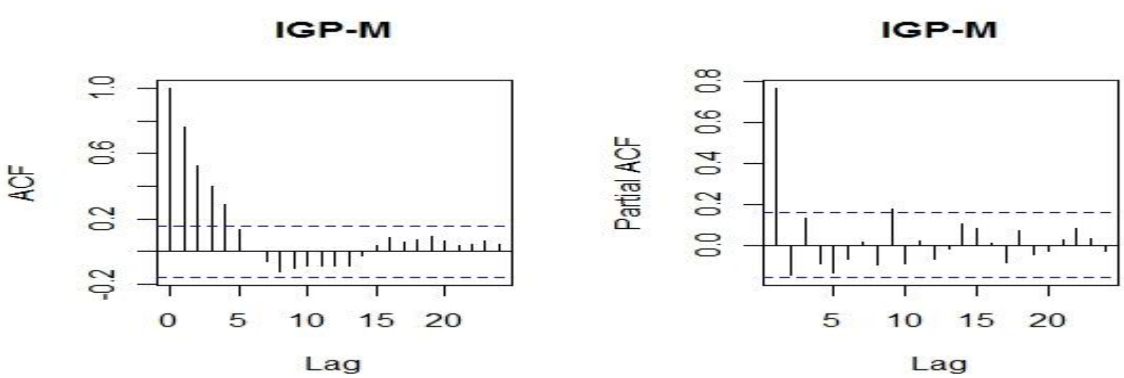

Figura 10 - Análise ACF/PACF

(Elaborado pelo Autor)

Pelos critérios de AIC, significância dos parâmetros e pelos resultados do Box-Test todos os horizontes de previsão foram modelados com AR(1).

Como no IPCA, para exemplificar a construção da previsão utilizando a modelagem Lasso, serão considerados 2 passos à frente. Todas as variáveis explicativas permanecem diferenciadas, e o procedimento para escolha de λ permanece por validação cruzada. Pela figura 11, vemos os dois log(λ) sugeridos, utilizaremos o que proporciona o erro médio mínimo (-6,8286).

Ao utilizarmos este λ, temos um modelo com intercepto e mais 4 variáveis explicativas, são referentes a dívida interna dos estados e municípios, dívida interna do governo federal e banco central, dívida total dos estados e municípios e M3 – fundos de renda fixa. Logo, temos que as variáveis de dívida interna dos estados e munícipios bem como a dívida interna do governo federal e banco central afetam tanto o IPCA quanto o IGP-M.

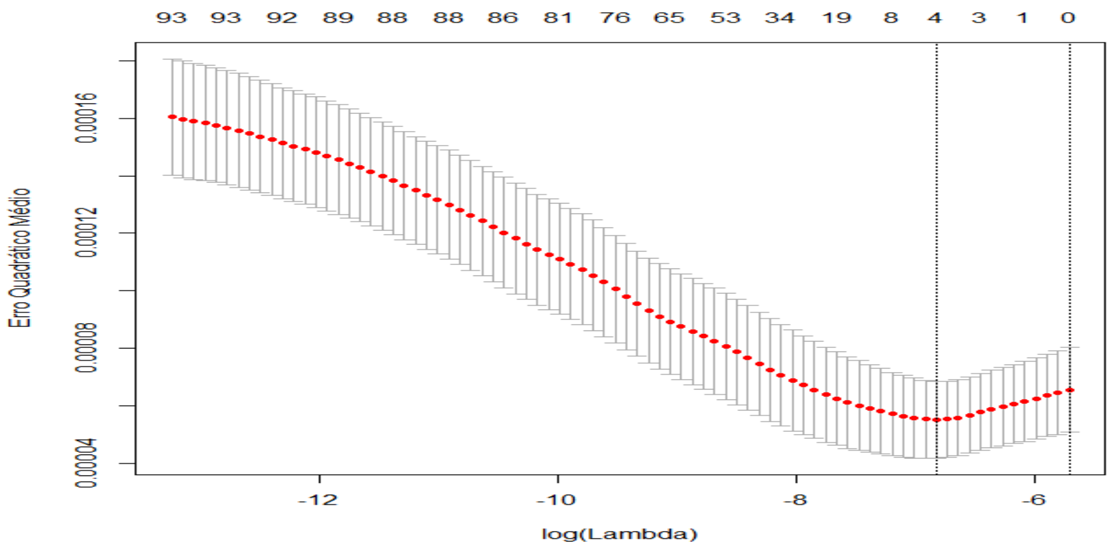

Figura 11 - Escolha λ

(Pacote glmnet – R Project)

No adaLasso, como no Lasso todas as variáveis explicativas foram diferenciadas, o λ foi utilizado o mesmo do procedimento anterior e a escolha de τ foi por validação cruzada. Temos um modelo com três variáveis explicativas, a variável de divida total dos estados e munícipios não permanece. Entretanto, este comportamento não se repete em todos os horizontes.

Através do WLadalasso, temos as mesmas três variáveis, exceto nos horizontes em que no adaLasso não há redução de variáveis, nestes casos, o WLadalasso que reduz a três variáveis. Neste modelo, dois passos a frente, o WLadalasso nos fornece o menor erro de previsão. Demonstrando que o acréscimo da defasagem é positivo para previsão.

Em geral, nos horizontes t+2 e t+3 a modelagem Lasso selecionou as 4 variáveis já mencionadas, e o adaLasso e WLadalasso reduziu para 3. Nos horizontes t+4, t+5, ..., t+12, a redução aconteceu somente ao modelarmos com WLadalasso. Quanto ao modelo escolhido para cada horizonte, podemos considerar as figuras 12 e 13 que mostram os erros quadráticos médios e erros médio absoluto das previsões, de t+2 até t+7 o WLadalasso fornece superioridade preditiva em relação a raiz do erro quadrático médio, exceto em t+4 onde o melhor desempenho quanto a este erro é da modelagem adaLasso. De t+8 até t+12 a modelagem AR(1) é superior. O erro médio absoluto tem comportamento semelhante ao RMSE.

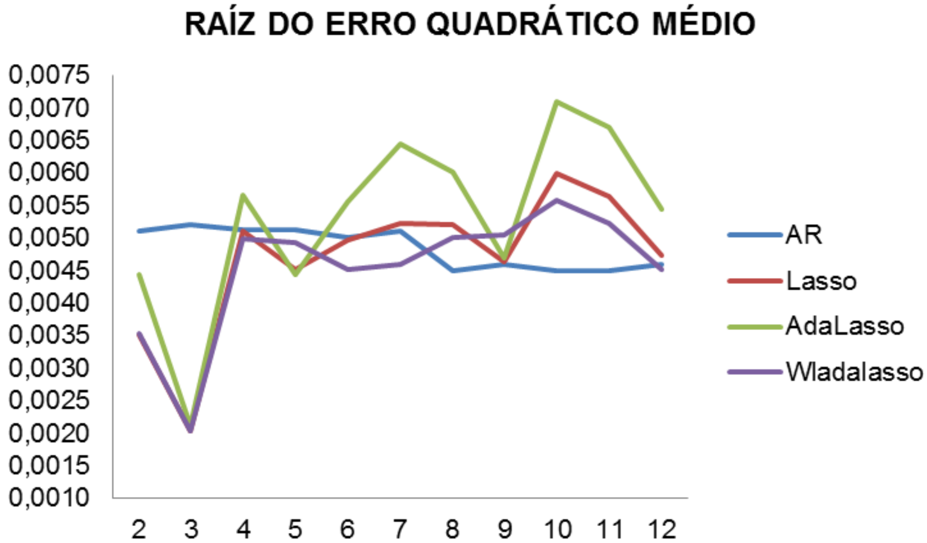

Figura 12- IGP-M RAIZ DO ERRO QUADRÁTICO MÉDIO

(Baseado em Medeiros et al. (2016))

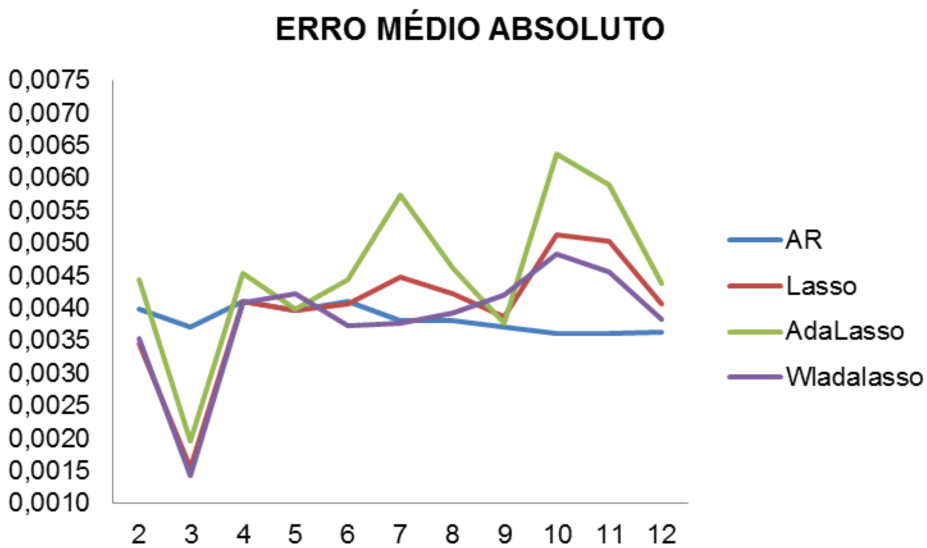

Figura 13- IGP-M ERRO MÉDIO ABSOLUTO

(Baseado em Medeiros et al. (2016))

A Tabela 2 apresenta o GW-test (Giacomini e White), bem como a Tabela 1. Temos alguns resultados significativos, diferente do caso IPCA, em que todas as previsões são estatisticamente iguais.

I. Em t+3 e t+7, a modelagem AR(1) é estatisticamente diferente de WLadalasso, e como os erros de previsão utilizando WLadalasso são menores, o WLadalasso apresenta superioridade significativa neste horizonte de previsão.

II. Em t+9, a modelagem WLadalasso é significativamente diferente de Lasso e adaLasso, entretanto seus erros de predição são maiores, logo, neste horizonte de previsão o WLadalasso não é a modelagem de melhor desempenho preditivo. Entre Lasso e Adalasso não há evidência estatística de diferença.

III. Em t+11 e t+12 o WLadalasso é significativamente diferente do Adalasso, apresentando menores erros de predição, sendo superior na comparação dos dois.

Como no IPCA, estes resultados evidenciam a contribuição da seleção de variáveis a modelagem da taxa de crescimento do IGP-M, mas não comprova significância estatística quanto à diferença de predição em todos os horizontes. Portanto, temos a solução do problema da alta dimensionalidade, mas não há comprovação estatística dos ganhos dessa utilização para todos os horizontes.

Em geral as variáveis selecionadas são: a dívida interna dos estados e municípios, dívida interna do governo federal e banco central e M3 – fundos de renda fixa.

a) **M3 – fundos de renda fixa:** tem uma carteira focada principalmente em títulos prefixados, o investidor sabe quando será resgatado.

		t+2					t+3		
Modelo	AR	Lasso	Adalasso	Wladalasso	Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,3983	0,4903	0,4535	AR	-	0,1056	0,4326	0,0456
Lasso	0,3983	-	0,8272	0,6992	Lasso	0,1056	-	0,7958	0,2350
Adalasso	0,4903	0,8272	-	0,3848	Adalasso	0,4326	0,7958	-	0,1909
Wladalasso	0,4535	0,6992	0,3848	-	Wladalasso	0,0456	0,2350	0,1909	-
		t+4					t+5		
Modelo	AR	Lasso	Adalasso	Wladalasso	Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,0925	0,1678	0,0869	AR	-	0,6857	0,7401	0,6619
Lasso	0,0925	-	0,9043	0,3281	Lasso	0,6857	-	0,6285	0,3462
Adalasso	0,1678	0,9043	-	0,0999	Adalasso	0,7401	0,6285	-	0,3488
Wladalasso	0,0869	0,3281	0,0999	-	Wladalasso	0,6619	0,3462	0,3488	-
		t+6					t+7		
Modelo	AR	Lasso	Adalasso	Wladalasso	Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,8062	0,7739	0,8438	AR	-	0,2370	0,6628	0,0076
Lasso	0,8062	-	0,5837	0,3258	Lasso	0,2370	-	0,8956	0,1123
Adalasso	0,7739	0,5837	-	0,3625	Adalasso	0,6628	0,8956	-	0,0962
Wladalasso	0,8438	0,3258	0,3625	-	Wladalasso	0,0076	0,1123	0,0962	-
		t+8					t+9		
Modelo	AR	Lasso	Adalasso	Wladalasso	Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9888	0,9828	0,9962	AR	-	0,9970	0,9978	0,9819
Lasso	0,9888	-	0,9161	0,1862	Lasso	0,9970	-	0,9867	0,0275
Adalasso	0,9828	0,9161	-	0,1145	Adalasso	0,9978	0,9867	-	0,0162
Wladalasso	0,9962	0,1862	0,1145	-	Wladalasso	0,9819	0,0275	0,0162	-
		t+10					t+11		
Modelo	AR	Lasso	Adalasso	Wladalasso	Modelo	AR	Lasso	Adalasso	Wladalasso
AR	-	0,9990	0,9990	0,9998	AR	-	0,9976	0,9987	0,9957
Lasso	0,9990	-	0,9963	0,0280	Lasso	0,9976	-	0,9952	0,1341
Adalasso	0,9990	0,9963	-	0,0084	Adalasso	0,9987	0,9952	-	0,0284
Wladalasso	0,9998	0,0280	0,0084	-	Wladalasso	0,9957	0,1341	0,0284	-
		t+12							
Modelo	AR	Lasso	Adalasso	Wladalasso					
AR	-	0,9481	0,9986	0,9883					
Lasso	0,9481	-	0,9980	0,5220					
Adalasso	0,9986	0,9980	-	0,0058					
Wladalasso	0,9883	0,5220	0,0058	-					

Tabela 2- Esta tabela mostra o p-valor do GW-test (a hipótese nula do GW é de que as previsões são estatisticamente iguais, (Baseado em Medeiros et al. (2016))

c. Curva Phillips

Ao não encontrarmos nenhuma variável relacionada a desemprego em nossas modelagens passamos a questionar se teríamos possivelmente encontrado uma variável mais robusta ou se como Medeiros et al. (2016) cita estas aplicações são evidencias contrárias a usual Curva de Phillips.

Considerando que as variáveis de taxa de crescimento de divida públicas foram importantes para previsão da taxa de crescimento dos dois índices, aplicamos os modelos de alta dimensão a ela, esperando encontrar um paralelo entre a Curva Phillips e nossos resultados.

Para previsão da variável taxa de crescimento da dívida pública foram selecionadas as seguintes variáveis: horas trabalhadas, resultado global do balanço de pagamentos, consumo aparente – bens de consumo, dívida fiscal do setor público, dívida interna dos estados e municípios, dívida total do governo federal e Banco Central, arrecadação das receitas federais, fluxo de caixa de previdência, índice de ações – Ibovespa, fundo de ações, dólar comercial, taxa de juros, M3 e M4. Frente a estas variáveis preditoras percebemos que a construção da variável taxa de crescimento de dívida pública passa por variáveis ligadas a emprego, tais como: horas trabalhadas, balanço de pagamentos e previdência.

Então, ao contrário de Medeiros et al.. (2016) não consideramos afirmar que este trabalho é uma evidencia contra a Curva de Phillips, uma vez que a construção das variáveis selecionadas nos modelos de alta dimensão passa por variáveis ligadas à relação de emprego e inflação. No entanto, ao invés de considerarmos apenas essa relação, podemos usar a variável de taxa de crescimento dívida pública numa função mais abrangente, uma vez que é construída com outras variáveis, igualmente importantes.

7. Considerações finais

O principal objetivo deste trabalho era avaliar a contribuição do WLadalasso na solução do problema de alta dimensionalidade e compará-lo com Lasso, adaLasso e modelos ARIMA em capacidade preditiva para os principais índices de inflação do Brasil, observando que informações mais recentes tendem a contribuir mais para as previsões, inspirado por Konzen e Ziegelmann (2016).

Os procedimentos de Lasso, adaLasso e WLadalasso selecionam variáveis explicativas que melhor explicam a variável resposta em uma estrutura dimensional elevada, no presente estudo eram 95 variáveis candidatas.

O WLadalasso foi o melhor modelo para previsão da inflação do índice IPCA do horizonte t+2 até t+6. Por outro lado, para horizontes maiores a modelagem AR(1) tem melhor desempenho. Entretanto, pelo teste de Giaomini e White para igualdade de acurácia preditiva, temos que as previsões de todos os modelos não são estatisticamente diferentes, o que pode ser resultado do tamanho da amostra. Em geral, as variáveis de petróleo e de dívida publica foram selecionadas para previsão.

Quanto a taxa de crescimento do IGP-M, de t+2 até t+8 o WLadalasso apresenta superioridade quanto aos demais modelos, com significância estatística nos horizontes t+3 e t+7, em relação a modelagem AR(1). Nos horizontes t+9 e t+10 a modelagem AR(1) domina, mas sem significância estatística. Para este índice, as variáveis de divida públicas também foram selecionadas na predição deste índice, com o acréscimo de uma variável monetária.

Temos alguns resultados principais: as variáveis de dívida pública representam a maioria dos preditores importantes para prever inflação a curto prazo; em Medeiros et al. (2016) comprovamos os ganhos da utilização do Lasso e adaLasso e neste trabalho aumentamos algumas precisões ao trabalharmos com a modelagem WLadalasso.

Por fim, os resultados encontrados não trouxeram nenhuma variável relacionada a emprego, o que é usual em termos econômicos. No entanto, ao investigar a variável de dívida pública como explicativa encontramos uma possível relação com emprego nas suas variáveis preditoras.

O presente estudo é especialmente importante por quantificar os resultados de diferentes métodos preditivos e de solução de problemas de alta dimensão. Além de auxiliar no entendimento das variáveis determinantes para previsão da inflação.

Referências

Giacomini, R. & White, H. Tests of conditional predictive ability. **Econometrica**, v. 74,1545-1578, 2016.

Medeiros, M. C. & Mendes, E. F. l1-estimation of high-dimensional timeseries models with exible innovations. **Working Paper**, 2015.

Medeiros, M. C. & Vasconcelos, G. F. R. (2015). Forecasting ination with highdimensional time-series models. **Working Paper**, 2015.

Tibshirani, R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. **Series B (Methodological)**, v. 58, 267-288, 1996.

Zhao, P. & Yu, B. On model selection consistency of Lasso. **The Journal of Machine Learning Research**, v. 7, 2541-2563, 2016.

Zou, H. The adaptive lasso and its oracle properties. **Journal of the American Statistical Association**, v. 101,1418-1429, 2006.

Konzen, E. & Zielgelmann F. LASSO-Type Penalties for Covariate Selection and Forecasting in Time Series, **Jornal of Forecasting**, 2016.

Medeiros, M. C. et al. Forecasting Brazilian Inflation with High Dimensional Models, **Working Paper**, 2016.

Sachsida, A. Inflação, Desemprego e Choques Cambiais: Uma Revisão da Literatura Sobre a Curva Philips no Brasil, **RBE**, v. 67,549-559, 2013.

Arruda, W. F. et al. Modelos Lineares e Não Lineares da Curva de Phillips para Previsão da Taxa de Inflação do Brasil, **RBE**, v. 64: 237-252, 2011.

Mendonça, H. F. & Santos, M. A. L. Credibilidade da Política Monetária e a Previsão do Trade-off entre inflação e Desemprego: Uma Aplicação para o Brasil, **Revista Economia,** 2006.

Konzen, Evandro. **Penalizações tipo Lasso na seleção de covariáveis em séries temporais**. 2014. 48. Mestre em Economia. UFRGS, Faculdade de Ciências Economicas.