



Modelos de previsão para uma série temporal da Dívida Interna Líquida do Setor Público como porcentagem do PIB

Andrea A. Prudente

Silvia Regina R. Lemos,

Valmir R. da Silva

Departamento de Estatística, CCEN, UFPE,

50740-540, Recife, PE

E-mail: deaprudente@gmail.com, silviareginalemos@ig.com.br,

estvalmir@hotmail.com

Resumo Este trabalho tem por objetivo avaliar determinados modelos de previsão aplicados a uma série temporal da Dívida Interna Líquida do Setor Público (DILSP) como porcentagem do Produto Interno Bruto (PIB) e realizar previsões. A fonte dos dados é a Seção de Finanças Públicas do Banco Central do Brasil. A série compreende o período de janeiro de 1991 até janeiro de 2007, tendo sido atualizada pela última vez em 28 de fevereiro de 2007. O objetivo principal do trabalho foi a determinação do comportamento da série (decompondo-a em fatores de tendência e sazonalidade) para a geração de previsões. Os modelos preditivos estudados foram os algoritmos de alisamento exponencial sazonal de Holt-Winters aditivo e multiplicativo, modelos sazonais auto-regressivos integrados de médias móveis e modelos sazonais auto-regressivos integrados de médias móveis com variável explicativa.

Palavras-chave Série temporal; métodos de previsão; alisamento exponencial; modelagem de Box-Jenkins; dívida interna líquida do setor público.

Introdução

Uma série temporal é uma coleção de observações de um fenômeno feitas sequencialmente ao longo do tempo. A característica mais importante deste tipo de dados é que observações vizinhas são dependentes e estamos interessados em analisar e modelar esta dependência. Enquanto em modelos de regressão com dados de corte transversal, por exemplo, a ordem das observações é irrelevante para a análise, em séries temporais a ordem dos dados é crucial.

A análise de uma série temporal é o processo de identificação das características, padrões e propriedades importantes da série, utilizados para descrever, em termos gerais, o seu fenômeno gerador. Dentre os diversos objetivos da análise de séries temporais, o maior deles é a previsão.

A previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisões, visando a determinados objetivos. Alguns modelos foram desenvolvidos com o objetivo de realizar previsões; a seleção

adequada destes modelos pode ser uma tarefa difícil dependendo dos modelos candidatos e das características em questão.

O incremento da complexidade e da satisfação estatística dos métodos de previsão de séries temporais não implica, necessariamente, melhora na acuidade da previsão. Métodos de previsão menos complexos normalmente permitem alcançar compreensão de suas suposições e limitações, e de interpretação de seus resultados. Assim, antes de se adotar um método de previsão mais complexo, é necessário avaliar os benefícios que um método dessa natureza pode gerar em relação ao custo de sua aplicação.

Este trabalho tem como objetivo fazer previsões através do algoritmo de Alisamento Exponencial Sazonal de Holt-Winters, SARIMA e SARIMAX para a série temporal da Dívida Interna Líquida do Setor Público (DILSP) (compreendendo o governo federal, Banco Central, governos estaduais, governos municipais e empresas estatais) como porcentagem do Produto Interno Bruto (PIB).

Para a determinação de qual modelo forneceu as melhores previsões foi utilizada como medida de precisão o erro quadrático médio. Foram consideradas previsões 1, 3, 6, 9 e 12 passos à frente a fim de avaliar as capacidades preditivas dos diferentes métodos em diferentes horizontes de previsão. Considerando a série temporal DILSP com n observações foram retiradas desta série as m últimas observações. Com essa nova série, sem as m últimas observações, foram geradas m previsões a partir do modelo estimado e obtidos os erros de previsão.

A plataforma computacional utilizada neste trabalho foi o programa R versão 2.4.1.

Dívida Interna Líquida do Setor Público

A dívida líquida corresponde ao saldo líquido do endividamento do setor público não-financeiro e do Banco Central com o sistema financeiro (público e privado), o setor privado não-financeiro e o resto do mundo. Entende-se por saldo líquido o balanceamento entre as dívidas e os créditos do se-

tor público não-financeiro e do Banco Central. O acompanhamento da dívida líquida do setor público vem sendo efetuado sistematicamente desde dezembro de 1990, como base para o cálculo das Necessidades de Financiamento do Setor Público. Dadas as várias alterações de padrão monetário ocorridas no período, a análise da série histórica em moeda nacional fica prejudicada. Tampouco pode ser feita a transformação da série para dólares norte-americanos, devido às várias mudanças na política cambial no período, que também distorcem a análise. Assim, é preferível apresentar a série como proporção do PIB.

Análise Exploratória

A série DILSP contém 193 observações mensais cobrindo o período de janeiro de 1991 a janeiro de 2007, tendo sido atualizada pela última vez em 28 de fevereiro de 2007. A fonte das suas observações foi a Seção de Finanças Públicas do Banco Central do Brasil.

Na Figura 1 temos as observações da série DILSP ao longo do tempo. A média amostral da série foi 33.46% do PIB com desvio padrão 11.88%. O menor valor observado foi de 13.10%, referente ao mês de novembro de 1991. O maior valor observado foi de 53.18%, que corresponde ao mês de janeiro de 2007. Notou-se uma leve assimetria negativa, com valor do coeficiente de assimetria igual a -0.21 e do coeficiente de curtose igual a -1.39 . Pode-se perceber que a série apresenta indícios de tendência, sazonalidade e não-estacionariedade.

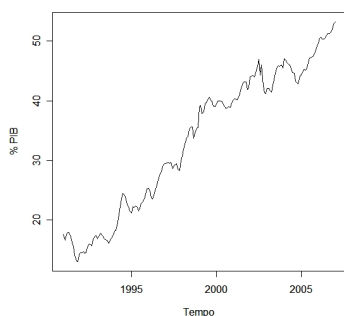


Figura 1: Série temporal da dívida interna líquida do setor público (DILSP) como porcentagem do PIB, no período de janeiro de 1991 a janeiro de 2007

A partir da decomposição da série foi comprovada a presença de sazonalidade e de tendência, ambas estocásticas. De acordo com o gráfico sazonal verificou-se a ocorrência de um pico nos meses de março e julho de cada ano. O gráfico da tendência mostrou que os dados apresentaram um crescimento estocástico. Por fim, o gráfico dos resíduos mostrou que julho e setembro de 2002 apresentaram

resíduos discrepantes quando comparados aos demais resíduos.

Foi analisado o correlograma da série e observou-se que as autocorrelações decrescem lentamente para zero indicando não-estacionariedade da série, conseqüentemente apresentando a necessidade de pelo menos uma diferenciação. Após verificação de não-estacionariedade da série fora aplicada a primeira diferenciação, a qual foi suficiente para torná-la estacionária. Segundo Morettin & Tolo (2006), de modo bastante geral, uma série temporal é estacionária se ela varia ao redor de uma média constante, com variância também constante.

Modelagem e Previsão

Para a modelagem da série DILSP, uma vez constatada a existência de tendência e sazonalidade, foram utilizados os seguintes métodos: algoritmo de alisamento exponencial de Holt-Winters (aditivo e multiplicativo), modelos SARIMA, modelos SARIMAX. Estes modelos foram comparados para verificação daquele que apresenta as melhores previsões. Todas as previsões serão comparadas mediante uma medida de precisão, erro quadrático médio. Serão consideradas previsões 1, 3, 6, 9 e 12 passos à frente a fim de avaliar as capacidades preditivas dos diferentes métodos em diferentes horizontes de previsão.

As previsões geradas pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente subestimaram os valores da série, com exceção das previsões obtidas pelo método sazonal multiplicativo referentes ao período de fevereiro a julho de 2006, que superestimaram os valores da série. Analisando o erro quadrático médio para previsões 1, 3, 6, 9 e 12 passos à frente para a série, observou-se que o método de Holt-Winters aditivo teve desempenho superior ao método sazonal multiplicativo, exceto para previsões 3 passos à frente, pois o erro quadrático médio foi menor para as previsões advindas do algoritmo aditivo.

A identificação do modelo SARIMA foi feita calculando-se os valores do critério BIC para vários modelos diferentes e escolhendo-se o modelo que apresentou o menor valor do critério. O critério BIC foi utilizado por ser consistente e parcimonioso [Morettin (2006)]. A função interna do programa R, `arma`, não retorna o valor do BIC; dessa forma, modificou-se essa função para que o valor do critério BIC fosse informado, criando-se a função `myarima`. Para esse procedimento de minimização dos valores do BIC utilizou-se uma função externa ao R, `bic`, que calcula o valor do BIC para vários modelos diferentes, associando valores para os ordens p e q de 0 a 3, d de 1, P e Q de 0 a 2 e D = 0 ou 1. Esses valores do BIC são armazenados em uma matriz, cuja posição (i, j) apresenta o valor do BIC referente ao modelo onde $p = i - 1$ e $q = j - 1$; ver Frery & Cribari-Neto (2005). Portanto, para a série DILSP foram obtidos 288 valores do BIC para os modelos consi-

derados. Desses 288 modelos, foram escolhidos dois, sendo eles SARIMA(0, 1, 0) \times (0, 1, 1) (SARIMA 1) e SARIMA(1, 1, 0) \times (0, 0, 0) (SARIMA 2) que apresentaram os menores BIC.

Após identificados os possíveis modelos que poderão representar adequadamente a série, foram estimados, usando o método de máxima verossimilhança, os parâmetros ϕ , θ , Φ e Θ , e foram produzidas previsões para a série com os dois modelos selecionados.

As previsões obtidas para a série através dos modelos SARIMA 12 passos à frente subestimaram os valores reais da série, exceto as previsões geradas pelo modelo SARIMA 1 referentes aos meses de maio, junho e julho de 2006 que foram muito próximas dos valores da série. Vale ressaltar que no modelo SARIMA 2 as previsões geradas de julho de 2006 a janeiro de 2007 foram todas iguais a 49.9101% do PIB. Analisando o erro quadrático médio para previsões 1, 3, 6, 9 e 12 passos à frente para a série, observou-se que o modelo SARIMA 1 apresentou melhor desempenho para previsões 3, 9 e 12 passos à frente, enquanto que o modelo SARIMA 2 apresentou uma melhor habilidade preditiva para previsões 1 e 6 passos à frente. Ao realizar a verificação de diagnóstico que consiste em constatar a qualidade do ajuste observou-se que o modelo estimado SARIMA 2 é adequado para representar os dados, pois os resíduos têm comportamento aproximado de ruído branco, segundo os gráficos dos resíduos padronizados e da função de autocorrelação dos resíduos, e não há evidência suficiente para rejeitar a hipótese de que o modelo SARIMA 2 está bem especificado, segundo o gráfico dos *p*-valores para o teste Ljung-Box [Ljung & Box (1978)].

Para o modelo SARIMAX foi utilizada uma tendência linear como regressor. A escolha do modelo SARIMAX que melhor representa a série da DILSP foi feita de forma semelhante à utilizada na escolha do modelo SARIMA e os dois modelos selecionados para a série, que apresentaram os menores BIC, foram SARIMAX(1,1,0) \times (0,0,0) (SARIMAX 1), SARIMAX(0,1,0) \times (0,1,1) (SARIMAX 2).

As previsões geradas pelo modelo SARIMAX 1 12 passos à frente referentes ao período de março a novembro de 2006 foram muito próximas dos valores da série, ao passo que as previsões dos outros meses subestimaram os valores realizados. Enquanto que as previsões produzidas pelo modelo SARIMAX 2 12 passos à frente referentes ao período de abril de 2006 a janeiro de 2007 superestimaram os valores da série, com exceção das previsões referentes aos meses de fevereiro e março de 2006 que subestimaram os valores da série. Vale ressaltar que a previsão para o mês de janeiro de 2007 pelo modelo SARIMAX 2 foi muito próxima do valor real da série. Analisando o erro quadrático médio para previsões 1, 3, 6, 9 e 12 passos à frente para a série, observou-se que o modelo SARIMAX 1 para as previsões 1, 3 e 12 passos à frente mostrou ter melhor capacidade preditiva, enquanto que o modelo SARIMAX 2 apresentou melhor desempenho para 6 e 9 passos

à frente. Ao realizar a verificação de diagnóstico que consiste em constatar a qualidade do ajuste observou-se que o modelo estimado SARIMAX 1 é adequado para representar os dados, pois os resíduos têm comportamento aproximado de ruído branco e não há evidência suficiente para rejeitar a hipótese de que o modelo SARIMAX 1 está bem especificado.

Conclusão

Conclui-se que o algoritmo de alisamento exponencial de Holt-Winters aditivo mostrou-se mais eficaz, no que diz respeito a capacidade preditiva, do que sua contrapartida multiplicativa. Para previsões de curto prazo (1 e 3 passos à frente), as previsões mais precisas foram aquelas obtidas a partir de modelos SARIMA. Por outro lado, para horizontes mais longos (6, 9 e 12 passos à frente), as previsões mais precisas foram aquelas advindas de modelos SARIMAX utilizando como variável explicativa uma tendência linear.

Referências

- [1] Frery, Alejandro C. & Cribari-Neto, F., “Elementos de Estatística Computacional Usando Plataformas de Software Livre/Gratuito”, Rio de Janeiro: IMPA, 2005.
- [2] Ljung, G. M. & Box, G. E. P., On a measure of lack of fit in time series models, *Biometrika*, 65 (1978) 297-303.
- [3] Morettin, Pedro A., “Econometria Financeira”, São Paulo: Associação Brasileira de Estatística, 2006.
- [4] Morettin, Pedro A. & Toloi, Clélia M. C., “Análise de Séries Temporais”, 2. ed. São Paulo: Edgard Blücher, 2006.