Higher Education
Australian Mathematical Sciences Institute

Pricing Contingent Claims on Cryptocurrencies

Omar Ghattas
Supervisor: Dale Roberts
The Australian National University

Australian Government
Department of Education

25/02/2015

1 Introduction

Cryptocurrencies, such as Bitcoin, have recently emerged as alternatives to fiat money for those seeking low transaction costs, anonymity and protection from the loose monetary policy of central banks. The growth of these new currencies can potentially be hindered by their high price volatility, negating the store of value characteristic that is desirable in any established currency. This will undoubtedly lead to the arrival of derivative securities (aka. contingent claims), in which these currencies are the underlying assets, in an attempt to manage that risk.

The traditional Black-Scholes option pricing approach, which revolutionised the role of mathematics in finance, has an obvious shortcoming when employed to pricing options on assets such as Bitcoins. The failure of the model is arguably a direct product of the related assumptions that there is an endless supply of the asset and the market is perfectly liquid (i.e., buying and selling does not change the market price). These assumptions are violated in the case of cryptocurrencies. First, coins are issued into circulation through the process of mining. They are rewarded to miners as an incentive for solving increasingly difficult and computationally intensive proof-of-work mathematical problems. The difficulty of these problems is directly related to the speed at which they are mined, which in turn affects the supply of coins. Second, as coins are in limited supply, the action of buying and selling (a necessary step in replicating and hedging a contingent claim) can drastically affect the market price of the asset.

1.1 Contingent Claims

Contingent claims are financial contracts whose payoff is dependent on the movement of an underlying asset. An example of a contingent claim is a European Call Option. These are contracts that provide the holder with the right, but not the obligation to buy an asset at a predetermined time in the future, for a specific price. The holder of this contract will only exercise his right if it is financially beneficial to do so, that is, the price of the asset underlying the option is cheaper than the price in the market. The famous Black-Scholes-Merton model provided a method by which these claims can be priced, however, the framework in which that model exists is often restrictive and unrealistic. In the case of Bitcoin, this is especially true as one of the main assumptions of the Black-Scholes model is that the underlying asset is in perfect supply, that is, a market participant is always able to buy as much of that asset as they please.

1.2 The Bitcoin Price Process

Cryptocurrencies are digital currencies that are decentralised. The most famous of these is Bitcoin, which has seen its popularity surge since its inception in 2009. However, Bitcoin prices are extremely volatile, and negate the store of value characteristic that is a requirement of credible currencies. That is, a deterrent in the way of more wide spread acceptance of digital currencies are the large and often severe fluctuations in their prices. These rapid movements can be characterised as *speculative bubbles*. Speculative bubbles are defined as being sudden, exponential increases in a series followed by a steep, almost vertical drop in the price to a more stable level. An examination of the Bitcoin to USD time series exhibits many of these phenomena. Thus, in order to price claims on Bitcoin, it is worthwhile to investigate modelling the bitcoin price, which would require an ability to model speculative bubbles.

2 The Model

A growing area in stochastic modelling is the modelling of Non-Causal Time Seres. This framework has been applied to a range of time series including macroeconomic data (Lanne, Raikkonen 2011), electronic currencies (Gourieroux, Hencic 2013) and index prices (Gourieroux, Zakoian 2013) in which speculative bubbles are present. This section will introduce the theoretical aspects of the model, and will be based heavily on the framework set out in Gourieroux, Jasiak 2014.

2.1 Causal Time Series

A time series, y_t is causal if:

$$x_t = \sum_{j=0}^{\infty} \psi_j e_{t-j}$$

where

$$e_t \sim \text{iid}(0, \sigma^2)$$

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

Intuitively, this means that the time series may be written as a linear combination of present and past white noise terms.

2.2 Causal-Noncausal Model

Causal-Noncausal model, or mixed models are stochastic processes for $y_t, t \in \mathbb{Z}$ defined by:

$$\Psi(B^{-1})\Phi(B)y_t = e_t$$

where:

$$B^k y_t = y_{t-k}$$

$$e_t \sim \text{iid}(0, \sigma^2)$$

$$\Psi(B^{-1}) = 1 - \psi_1 B^{-1} - \psi_2 B^{-2} - ... - \psi_s B^{-s}$$

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_r B^r$$

In order for stationarity to hold, the roots of both polynomials must lie outside the unit circle. Clearly, when $\psi_1 = ..\psi_s = 0$, then y_t is the purely causal process as

defined in section 2.1.
With these assumptions in place, a unique stationary solution exists, given by:

$$y_t = \sum_{j=-\infty}^{\infty} \zeta_j e_{t-j}$$

In contrast to the causal model, this representation means that the time series may be written as a linear combination of present, past and future white noise terms.

An important note to consider is that when errors are assumed to be normally distributed, the model is not identifiable as we are not able to distinguish between causal and non-causal components. Due to this, I follow the method of (Gourieroux, Hencic 2013) and use cauchy distributed errors. Recall that:

$$X \sim \text{Cauchy}(0, \gamma) \implies f_X(x) = \frac{1}{\pi}\left[\frac{\gamma}{x^2 + \gamma^2}\right]$$

2.3 Economic Interpretation

A natural question of this model is why the model depends on expected future errors. First, the model can be seen as a solution to the Rational Expectations Hypothesis in economics. The idea of this is that economic agents make decisions regarding the future value of economic variables based on their past experiences, available information as well as their rational outlook. Non causal models provide a method to deal with the fact that the econometricians information set is smaller than the agents information set. Thus, non causal models allow for taking effect of the true information set without explicitly specifying it. (Hansen, Sargent 1991). Thus, in non causal models, the predictable parts of future errors have an effect on our expectations. (Lane, Saikkonen 2008).

2.4 Estimation

In order to fit the non causal model, we can use a modified version of the traditional Box Jenkins approach to fitting causal time series models.

2.4.1 Model Selection

The total autoregressive order, $p = r + s$ may be inferred from the autocorrelation function (ACF) (Andrews, Davis 2009). However, the values of r, s cannot be inferred prior to fitting the models to the data. Further, the confidence intervals of the ACF are no longer relevant due to the fact that autocorrelations do not theoretically exist for a distribution with infinite variance. I use the methodology outlined in Hencic, Gourieroux 2013 to find the modified confidence interval in dealing with the data.

2.4.2 Parameter Estimation

An Approximate Maximum Likelihood estimate is required here (Lanne, Raikkonen 2008), which is similar to the traditional MLE estimator but ignores observations before time r, and following time $T-s$ where T is the total number of observations in our sample. Intuitively, since our MLE estimate now depends on future observations, it is impossible to sum over all observations as we would always require an extra observation to complete that task. Mathematically, this amounts to finding the parameter values that maximise the log-likelihood function:

$$(\hat{\Psi}, \hat{\Phi}, \hat{\gamma}) = \text{argmax}_{\Psi,\Phi,\gamma} \sum_{t=r+1}^{T-s} \log[g[\Psi(B^{-1})\Phi(B)y_t]; \gamma]$$

Where $g[., \gamma]$ denotes the probability density function of e_t with scale parameter γ.

3 Analysis

The application of this framework to Bitcoin samples was carried out in Hencic, Gourieroux 2013 for a sample of 150 observations. I hope to extend their analysis by applying the model to a sample of 295 observations spanning from 01/04/2014 until 20/01/2015.

3.1 The Data

We look at 295 observations of the Bitcoin/USD exchange rate over the period 01/04/2014 - 20/01/2015. The data contain many instances that could be characterised as being speculative bubbles.

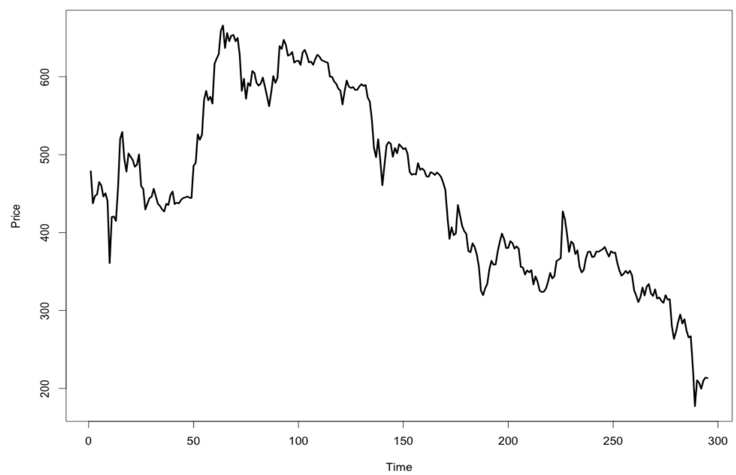

Figure 1: Bitcoin Price Process

3.1.1 Autocorrelation Function Analysis

For a stationary process, the autocorrelation is defined as:

$$\rho_k = \text{Correlation}(X_t, X_{t+k}) = \frac{\gamma_k}{\gamma_0} = \frac{\text{Covariance}(X_t, X_{t+k})}{\text{Variance}(X_t)}$$

A plot of the Autocorrelation function is known as a Corellogram, and is useful in inferring the required number of lags to be included in a time series model.

Following the methodology of Hencic, Gourieroux 2013, the confidence intervals for the Cauchy distributed errors must be approximated by simulation. I draw e_1, e_2, e_3 from the standard normal distribution and build the ratio $Z = \frac{e_1 e_3^2}{e_2}$. From a sample of 1000 values of Z, multiplied by $\frac{ln(T)}{T}$, I obtain a confidence interval of [-0.2363,0.3501]. Plotting the ACF below shows that the series is non-stationary due to a slowly decaying ACF. The modified confidence interval is plotted in red.

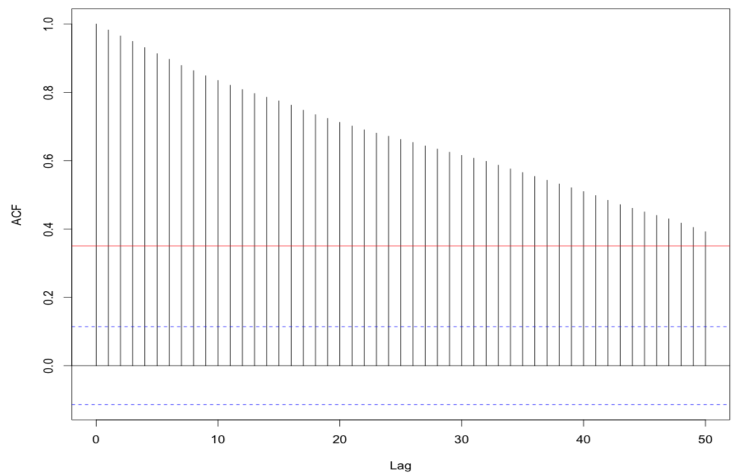

Figure 2: Autocorrelation of Bitcoin Price

This indicates that the data need to be de-trended in order to achieve stationarity.

3.1.2 Decomposition of Time Series

In analysing time series, we can consider the following model:

$$y_t = \text{seasonal}_t + \text{trend}_t + e_t$$

That is, the observation at time t is composed of both a seasonal and trend component, as well as some error term. In the Bitcoin case, it would be hard to justify the existence of any seasonal component, and so a more appropriate model is:

$$y_t = \text{trend}_t + e_t$$

Using a Natural Cubic Splines smoother, the trend component of the time series is identified. Cubic splines estimates a function by the use of a piecewise cubic polynomial with some added constraints. These constraints allow for continuity between where the individual piece wise cubics are fitted, as well as continuous first and second derivatives at those segments to allow for a smoother fit. The smoothed trend is depicted in red below:

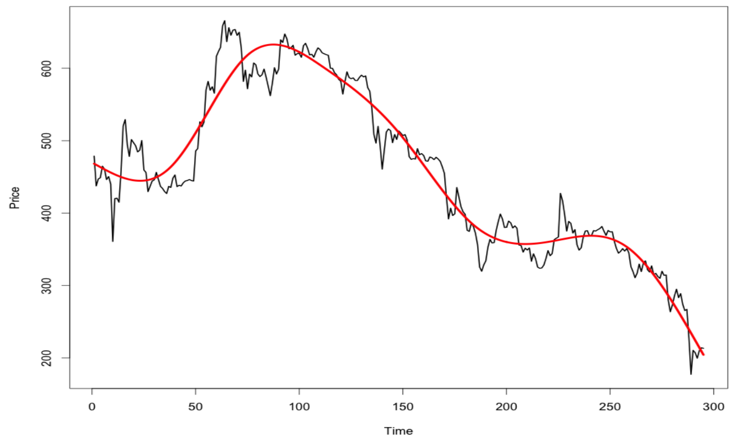

Figure 3: Trend Decomposition

The fluctuation of Bitcoin price around the trend component is the error or residual component of the time series. The de-trended series, is the time series less the trend, and is depicted below:

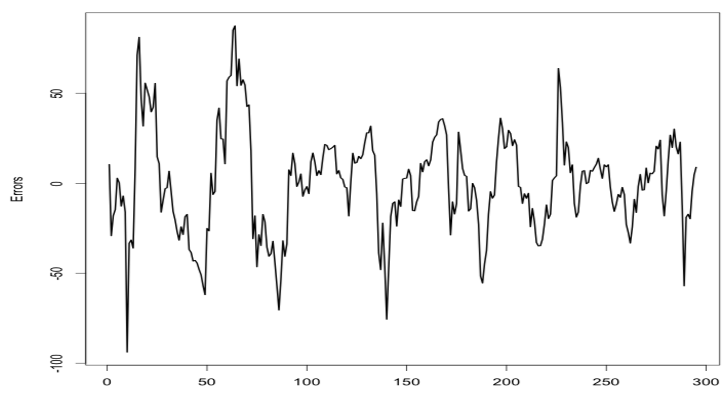

Figure 4: De-trended Series

This series is now stationary, and will hopefully admit itself to be analysed under the developed framework.

3.1.3 ACF Analysis of De-Trended Series

We again analyse the correllogram of the de-trended series, and this is plotted below:

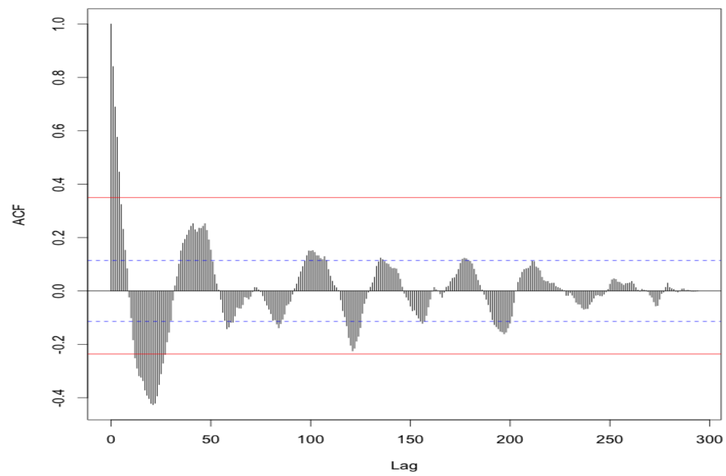

Figure 5: Autocorrelation of De-Trended Series

This is an interesting result, the correlogram fluctuates over the lags while tapering off to zero as the lag value becomes larger. This implies an Autoregressive model in traditional time series fitting, however, in our model, this would indicate a Mixed Autoregressive model (MAR). The term mixed refers to the fact that we have both forward and backward (causal/non-causal) components. Due to the structure of the ACF in this case, we are not able to infer the order of the MAR model, and must consider the Partial Auto Correlation Function (PACF) instead to achieve this.

The PACF is defined as the conditional correlation of X_{t+k} with X_t given $X_{t+1}, ..., X_{t+k-1}$. In this case, the PACF indicates that the order of our MAR(r, s) model is $p = r + s = 3$. Ignoring the trivial cases of MAR$(0, 3)$, MAR$(3, 0)$, this indicates that the best model to fit is either a MAR$(1, 2)$, or a MAR$(2, 1)$.

3.2 Model Fitting

In this section I use the results in the previous section to fit the MAR$(1,2)$ and MAR$(2,1)$ models. After fitting these models and using the Approximate MLE outlined in 2.4, I will assess the suitability of the models by considering their standard errors.

3.2.1 MAR(r=1,s=2)

Recall that the MAR(r,s) model, as outlined in 2.2, is given by:

$$\Psi(B^{-1})\Phi(B)y_t = e_t$$

Which in this case, expands to:

$$(1-\psi_1 B^{-1} - \psi_2 B^{-2})(1-\phi_1 B)y_t = e_t$$

$$\implies (1-\phi_1 B - \psi_1 B^{-1} + \psi_1\phi_1 - \psi_2 B^{-2} + \psi_2\phi_1 B^{-1})y_t = e_t$$

$$\implies y_t - \phi_1 y_{t-1} - \psi_1 y_{t+1} + \psi_1\phi_1 y_t - \psi_2 y_{t+2} + \psi_2\phi_1 y_{t+1} = e_t$$

$$\implies y_t(1+\psi_1\phi_1) + y_{t+1}(\phi_1\psi_2 - \psi_1) - \psi_2 y_{t+2} - \phi_1 y_{t-1} = e_t$$

Plugging this expression into the approximate MLE:

$$(\hat{\Psi},\hat{\Phi},\hat{\gamma}) = \text{argmax}_{\Psi,\Phi,\gamma} \sum_{t=1+1}^{295-2} \log[y_t(1+\psi_1\phi_1) + y_{t+1}(\phi_1\psi_2 - \psi_1) - \psi_2 y_{t+2} - \phi_1 y_{t-1};\gamma]$$

I use R to solve this numerically by minimising the negative of the log likelihood, see Appendix A for the code used. Further, I use the fact that the diagonals of the inverse hessian matrix are the variances of the parameters. I take the square root of these values to find the standard error of each parameter. Further, recall that the t-ratio is defined as the parameter estimate divided by its standard error. With large enough samples, t-ratios greater than 1.96 in absolute value indicate a statistically significant estimate at the 95% confidence level. The results are tabulated below:

Table 1: Results of Approximate MLE of MAR(1,2)

Parameter	Estimate	Standard Error	t-ratio
ψ_1	-0.4628	0.0255	-18.1355
ψ_2	-0.0628	0.0278	-2.2577
ϕ_1	0.3994	0.0223	17.9075
γ	4.2042	0.3267	12.8702

Therefore, the model is:

$$(1 + 0.4628B^{-1} + 0.0628B^{-2})(1 - 0.3994B)y_t = e_t$$

The non causal polynomial has complex values roots that lie outside the unit circle. The causal polynomial has a real valued root that also lies outside the unit circle. Therefore, both polynomials satisfy the model conditions and admit a two sided moving average representation. Further, the t-ratios of all parameter estimates seem to be significant at the 95% confidence level. This model seems to be a good fit to the data, however, it is still worthwhile to fit the other possible MAR model.

3.2.2 MAR(r=2,s=1)

Using an identical method to 3.2.1:

$$(1 - \psi_1 B^{-1})(1 - \phi_1 B - \phi_2 B^2)y_t = e_t$$
$$\implies y_t(1 + \psi_1\phi_1) + y_{t-1}(\psi_1\phi_2 - \phi_1) - \phi_2 y_{t-2} - \psi_1 y_{t+1} = e_t$$

and thus:

$$(\hat{\Psi}, \hat{\Phi}, \hat{\gamma}) = \text{argmax}_{\Psi,\Phi,\gamma} \sum_{t=2+1}^{295-1} \log[y_t(1 + \psi_1\phi_1) + y_{t-1}(\psi_1\phi_2 - \phi_1) - \phi_2 y_{t-2} - \psi_1 y_{t+1}; \gamma]$$

Table 2: Results of Approximate MLE of MAR(2,1)

Parameter	Estimate	Standard Error	t-ratio
ψ_1	-0.0723	0.0567	-1.2748
ϕ_1	0.9910	0.0425	23.3038
ϕ_2	-0.0805	0.0400	-2.0122
γ	6.2001	0.4970	12.4743

Therefore, the model is:

$$(1 + 0.0723B^{-1})(1 - 0.9910B + 0.0805B^2)y_t = e_t$$

In this case, the non causal polynomial has a real valued root that lies outside the unit circle. The causal polynomial has real valued roots that are both outside the unit circle. This model too admits an infinite autoregressive representation. However, the t-ratio for ψ_1 is smaller than 1.96 in absolute value, indicating that this parameter is not significant at the 95% level. This would mean a purely non-causal model may be better than the MAR (2,1), which defeats the purpose of using this novel method to model speculative bubbles. For this reason, the MAR(1,2) model is a more appropriate representation of the bitcoin price.

Proceeding with the finalised model, we can graphically depict how well our model adheres to the detrended bitcoin price below:

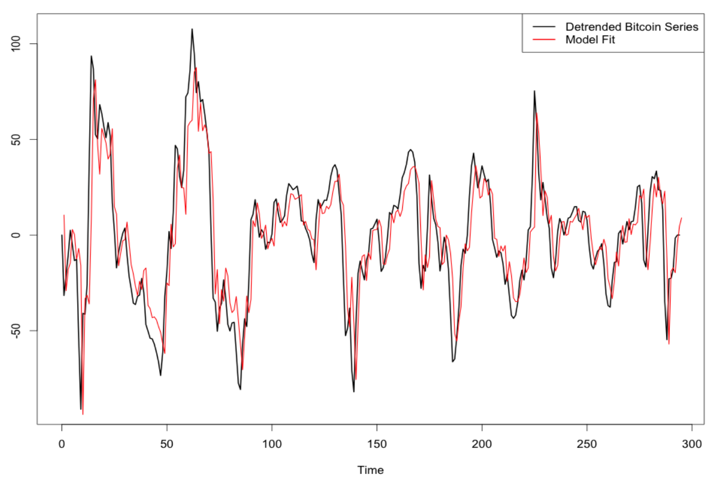

Figure 6: MAR(1,2) Model Fit

The fitted model seems a suitable fit to the data, it is interesting to note that the non-causal parameters, ψ_1, ψ_2 control the rate of increase of the bubble, and the causal parameter ϕ_1 controls the rate of descent of the bubble. That is, the non causal parameters dictate the growth phase of the bubble, while the causal parameters control the decline of the bubble.

4 Conclusion

The aim of this project was to investigate the pricing of contingent claims on cryptocurrencies such as Bitcoin. We approached this task by first attempting to model volatility of Bitcoin prices by making use of the novel and growing concept of non-causal time series models that would be able to capture the speculative price bubbles evident in the price. We applied the methodology outlined in Hencic, Gourieroux 2013 to a larger data set spanning from 01/04/2014 - 20/01/2015. The final model fit was a MAR$(2, 1)$. The next steps of this project are to be able to make predictions of the Bitcoin price based on our model as well as connect the pricing of Bitcoin to the pricing of contingent claims on Bitcoin.

5 Acknowledgements

I would like to thank my supervisor, Dr Dale Roberts for both his patience and guidance throughout this project. I would also like to thank AMSI for giving me first hand exposure to a research environment, this has undoubtedly left me in a more confident position as I enter into my Honours year at the ANU.

6 Appendix

6.1 R Code

```
#define g(x;gamma)
g  <-  function(x,gamma) {
  (1/pi) * ( gamma / (x^2 + gamma^2) )
}

# vector of detrended bitcoin series
y <- as.vector(detrend)

T=length(detrend)
r=2
s=1

# the function to be optimised is optimised over the first argument.
fun  <- function(x){
  phi = x[1:2]
  psi = x[3:4]
  gamma = x[5]

 ## AMLE OF MAR ( r = 2 , s = 1 )
  # PERFORM SUM
  res  <-  0.0
  for(t in seq(r + 1,T-s)){
    # calculate A
    A  <-  (y[t]   * (psi[1]*phi[1] + 1)
            + (psi[1]*phi[2] + phi[1]) * y[t-1]
            - phi[2]* y[t-2]
            - psi[1] * y[t+1])
    # increment sum
    res  <-  res + log(g(A,gamma))

  }
  #nlm performs the argmin, so return -out
  return(-res)
}
```

```
# Optimization
out <- nlm(fun,
           c(0.5,0.5,# phi
             0.5,0,# psi
             0.5 ),hessian=TRUE)# gamma

phi_hat  <- out$estimate[1:2]
psi_hat  <- out$estimate[3:4]
gamma_hat <- out$estimate[5]

## AMLE OF MAR ( r = 1 , s = 2 )
  # PERFROM SUM
  res <- 0.0
  for(t in seq(r + 1,T-s)){
    # calculate A
    A  <-  (y[t]  * (psi[1]*phi[1] + 1)
           + (phi[1]*psi[2] -psi[1]) * y[t+1]
           - psi[2]* y[t+2]
           - phi[1] * y[t-1])
    # increment the sum
    res  <-  res + log(g(A,gamma))

    #cat(t, " ", A, " ", gamma, " ", log(g(A,gamma)), " ", res, "\n")
  }
  #nlm performs the argmin, so return -out
  return(-res)
}

# Optimization
out <- nlm(fun,
           c(0.5,0,# phi
             0.5,0.5,# psi
             0.5 ),hessian=TRUE)# gamma
```

References

[1] Lanne, M., Saikkonen, Noncausal vector autoregression. Econometric Theory, 29(03), 447-481.

[2] Lanne, M., Saikkonen, P. (2011). Noncausal autoregressions for economic time series. Journal of Time Series Econometrics, 3(3).

[3] Lanne, Markku, and Pentti Saikkonen. "Modeling expectations with noncausal autoregressions." Available at SSRN 1210122 (2008).

[4] Hencic, A., Gouriroux, C. (2015). Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rates. In Econometrics of Risk (pp. 17-40). Springer International Publishing.

[5] Gouriroux, C., Zakoian, J. M. (2013). Explosive Bubble Modelling by Noncausal Process (No. 2013-04).

[6] Gourieroux, C., Jasiak, J. (2014). Filtering, Prediction and Estimation of Noncausal Processes.

[7] Hansen, L. P., Sargent, T. J., Heaton, J., Marcet, A. (1991). Rational expectations econometrics (pp. 45-76). Boulder, CO: Westview Press.

[8] Andrews, B., Calder, M., Davis, R. A. (2009). Maximum likelihood estimation for α-stable autoregressive processes. The Annals of Statistics, 37(4), 1946-1982.