Fitting autoregressive models by reversible jump Markov chain Monte Carlo

Jared McBride

Applied Mathematics
University of Arizona

May 10, 2022

Outline

Introduction:

Why fit an autoregressive model?

- Approximate the power spectrum of processes
- AR fitting
- Whitening filter and LPC problem
- Stepping stone to full ARMA fitting and rational approximation of spectra
- Order Selection?

Autoregressive models of order p, $AR(p)$ have the form

$$Y_n + a_1 Y_{n-1} + \cdots + a_p Y_{n-p} = \mu + \varepsilon_n \qquad \varepsilon_n \sim N(0, \sigma^2) \quad \text{i.i.d.}$$

If $Y = (Y_n,\ n > -\infty)$, $\qquad \varepsilon = (\varepsilon_n,\ n > -\infty)$, $\qquad a = (1, a_1, a_2, \ldots, a_p)$
Then (centering Y)

$$(a \star Y)_n = \varepsilon_n$$

It can be shown that (letting $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_p z^{-p}$)

$$S_Y(z) = \frac{\sigma^2}{A(z)A^*(z^{-*})} = \frac{\sigma^2}{(1 - z_1 z^{-1})(1 - z_1^* z) \cdots (1 - z_p^{-1})(1 - z_p^* z)}$$

For Y to be stationary it is required that $z_1, \ldots, z_p \in \mathbb{D}$

Given a model and data

- θ, model parameters
- y, data or observation

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta)$$

In our case given data y and parameters θ,

$$y = (y_n, n = 1, 2, \dots, M), \qquad \theta = (p, \sigma^2, z_1, \dots, z_p) = (p, v, z)$$

We will write

$$\pi(p, v, z|y) \propto \pi(y|p, v, z)\pi(v, z|p)\pi(p)$$

Maximal likelihood use optimization to find the a maximum.

Likelihood of AR model

Observe that since $\varepsilon \sim CN(0, \sigma^2 I_M)$ is Gaussian, $\tilde{Y} = (Y_n,\ n = 1, \dots, M)$ is also Gaussian. So, let $\mathrm{var}(\tilde{Y}) = \sigma^2 Q_M$ Then the density of $\tilde{Y}$ my be written

$$f_{\tilde{Y}}(y) = (\pi\sigma_\varepsilon^2)^{-M} |Q_M|^{-1} \exp\left(\frac{-1}{\sigma_\varepsilon^2} y^* Q_M^{-1} y\right)$$

or rather

$$\pi(y|p, v, z) = (\pi v)^{-M} |Q_M(z)|^{-1} \exp\left(\frac{-1}{v} y^* [Q_M(z)]^{-1} y\right)$$

We are mainly concerned with the following ratios

$$\frac{\pi(y|p, v, z')}{\pi(y|p, v, z)} = \frac{|Q_M(z)|}{|Q_M(z')|} \exp\left(\frac{-1}{v}\left(y^*[Q_M(z')]^{-1}y - y^*[Q_M(z)]^{-1}y\right)\right)$$

$$\frac{\pi(y|p, v', z)}{\pi(y|p, v, z)} = \left(\frac{v}{v'}\right)^M \exp\left(\left(\frac{-1}{v'} + \frac{1}{v}\right) y^*[Q_M(z)]^{-1} y\right)$$

A brief intro to RJMCMC

A generalization of Metropolis-Hastings,
Recall for (regular) MH a purposed move from x to x' is accepted with probability

$$\alpha = 1 \wedge \frac{\pi(x')q(x,x')}{\pi(x)q(x',x)}$$

For RJMCMC we wish to include more general spaces

$$C = \bigcup_{p=1}^{p_{\max}} C_k \qquad \text{where } C_k = \{k\} \times \mathbb{R}^+ \times \mathbb{R}^k$$

Dimension matching how to jump from C_k to $C_{k'}$ in a way will make sense and allow for detailed balance.

$$g : C_k \times \Omega_r \to C_{k'} \times \Omega_{r'}$$

So that

- $k + r = k' + r'$
- g is a diffeomorphism

A brief intro to RJMCMC

$$\begin{aligned}
\alpha &= 1 \wedge \frac{\pi(p', v', z'|y)}{\pi(p, v, z|y)} \cdot \frac{q(v, z|v', z', p, p')q(p|p')}{q(v', z'|v, z, p, p')q(p'|p)} \left|\frac{\partial(v', z')}{\partial(v', z')}\right| \\
&= 1 \wedge \frac{\pi(y|v', z', p')\pi(v', z'|p')\pi(p')}{\pi(y|v, z, p)\pi(v, z|p)\pi(p)} \cdot \frac{q(v, z|v', z', p, p')q(p|p')}{q(v', z'|v, z, p, p')q(p'|p)} \left|\frac{\partial(v', z')}{\partial(v', z')}\right| \\
&= 1 \wedge \frac{\pi(y|v', z', p')}{\pi(y|v, z, p)} \cdot \frac{\pi(v', z'|p')\pi(p')}{\pi(v, z|p)\pi(p)} \cdot \frac{q(v, z|v', z', p, p')q(p|p')}{q(v', z'|v, z, p, p')q(p'|p)} \left|\frac{\partial(v', z')}{\partial(v', z')}\right|
\end{aligned}$$

$$\alpha = 1 \wedge \text{(likelihood ratio)} \times \text{(prior ratio)} \times \text{(proposal ratio)} \times \text{(Jacobian)}$$

The parameters of the AR models will be specified by

- the variance of the white noise process
- the poles of the transfer function

The parameter space we wish to explore is

$$C = \bigcup_{p=1}^{p_{\max}} C_k \qquad \text{where } C_k = \{k\} \times \mathbb{R}^+ \times \mathbb{D}^k$$

This many of the modeling choices in the sequel follow those of Green in [**?**]. Let

$$X = (P, V, Z_1, Z_2, \ldots, Z_P) \in C$$

Priors

- The order P will be a Poisson distribution conditioned on $P \leq p_{\text{max}}$

$$f_P(p) = \frac{\lambda^p e^{-\lambda}}{p!} \left(\sum_{j=0}^{p_{\text{max}}} \frac{\lambda^j e^{-\lambda}}{j!} \right)^{-1}$$

- For the variance V let $u \sim \text{Unif}([-\hat{\beta}, \hat{\beta}])$ and

$$V = e^u \qquad \text{so that} \qquad f_V(v) = \frac{1}{2\hat{\beta}v} \mathbf{1}_{[-\hat{\beta},\hat{\beta}]}(\log(v))$$

- The p poles are *independent* and chosen randomly (uniformly) over $\mathbb{D}$,

$$f_{Z_j}(z_j) = \frac{1}{\pi} \mathbf{1}_{\mathbb{D}}(z_j) \qquad \text{for } j = 1, \ldots, p$$

There are three move types:

- Change in variance, with probability β_p (p is the number of poles)
- Change in pole position, with probability π_p
- Birth or death of pole, with probabilities b_p, and d_p, respectively. More precisely, a move from
  - $mathcalC_p$ to $mathcalC_{p+1}$ occurs with probability b_p
  - $mathcalC_{p+1}$ to $mathcalC_p$ occurs with probability d_p

The probabilities observe the following

- $\beta_p + \pi_p + b_p + d_p = 1$ for all p
- $d_0 = \pi_0 = b_{p_{\max}} = 0$
- $b_p = c \min\{1, f_P(p+1)/f_P(p)\}$, and
  $d_p = c \min\{1, f_P(p)/f_P(p+1)\}$
  for some c as large as possible so that $b_p + d_p \leq 0.9$

Proposals and acceptance ratios

Change in variance. (Move within C_p)

Program in Applied Mathematics

The new variance V' will be be so that

$$\log(V'/V) \sim \text{Unif}([-\hat{\beta}, \hat{\beta}])$$

meaning $V' = Ve^u$ where $u \sim \text{Unif}([-\hat{\beta}, \hat{\beta}])$ and so,

$$f_{V'|V}(v', v) = \begin{cases} \frac{1}{2\hat{\beta}v'}, & v' \in \left[ve^{-\hat{\beta}}, ve^{\hat{\beta}}\right] \\ 0, & \text{otherwise} \end{cases}$$

Acceptance ratio reduces to

$$\alpha = 1 \wedge (\text{likelihood ratio})$$

Proposals and acceptance ratios

Change in pole position. (Move within C_p)

Randomly select $j = 1, \ldots, p$ (uniformly) and the pole Z_j will be perturbed by $\tilde{u} \sim CN(Z_j, \hat{\pi})$ conditioned on $Z_j' = Z_j + \tilde{u} \in \mathbb{D}$. This gives

$$f_{Z_j'|Z_j}(z', z) \propto \begin{cases} \frac{1}{\pi\hat{\pi}} \exp\left(\frac{-1}{\hat{\pi}}|z' - z|^2\right) [I(z)]^{-1}, & z' \in \mathbb{D} \\ 0, & \text{otherwise} \end{cases}$$

where $I(z) = \dfrac{1}{\pi\hat{\pi}} \displaystyle\int_{\mathbb{D}} \exp\left(\frac{-1}{\hat{\pi}}|z' - z|^2\right) dz'$ w.r.t Lebesgue measure on $\mathbb{C}$.

Acceptance ratio reduces to

$$\alpha = 1 \wedge (\text{likelihood ratio }) \times \frac{I(z')}{I(z)}$$

- Birth of a pole. Move from C_p to C_{p+1} All we do is append a pole drawn from the uniform distribution on the unit disk.

$$\alpha = 1 \wedge (\text{likelihood ratio}) \times \frac{\pi^p \lambda}{p+1} \frac{d_{p+1}}{b_p}$$

- Death of a pole. Move from C_p to C_{p+1} All we do is delete a pole drawn from the uniform distribution of current poles.

$$\alpha = 1 \wedge (\text{likelihood ratio}) \times \frac{p+1}{\pi^p \lambda} \frac{b_p}{d_{p+1}}$$

Results?

- low acceptance rate $\approx$ 0.0005 (after N = 50,000 steps)
- Instability $\alpha \longrightarrow -\infty$

Thank you!