Non-Normal Time Series Estimation

Replication Study: Theory

Introduction

The general structure of the autoregressive gamma model explored in this paper is given as follows:

$$X_n = \alpha * X_{n-1} + \varepsilon_n$$

Where $\varepsilon_n \overset{\text{iid}}{\sim} \text{Gamma}(\alpha, \nu)$, for shape parameter α and scale parameter ν.

While α informs the magnitude of the autoregressive process, it is not the same as an AR parameter in a standard Gaussian system. Indeed,

$$\alpha * X = \sum_{i=1}^{N(X)} W_i$$

Where $W_i \overset{\text{iid}}{\sim} \text{Exponential}(\alpha)$ and $N(X) \mid X = x \sim \text{Poisson}(\lambda = p\alpha)$ for $p \in [0, 1)$ Thus, it is p that informs the rate of the compound Poisson process, and thus the magnitude of time-dependency in the system.

As we will explore later, p functions similarly to the AR parameter in the Gaussian system. The shape parameter α has an impact both in the shape of the gamma distribution and the consequent autoregressive structure, while ν solely informs the structure of the underlying Gamma distribution.

Proof of 2.1.1

In this case, what is crucial is that for a fixed X, the poisson pmf is still multiplied by x since it is a poisson Process, i.e.

$$\mathbb{P}[N(X) = n \mid X = X] = \frac{(\lambda x)^n}{n!} \exp(-\lambda x)$$

From the above, the Laplace transform of $\alpha * X$ for a fixed $X = x$ can be computed by recalling that the sum of independent random variables is equal to the convolution of their probability distributions. (source)

$$
\begin{aligned}
\{\mathcal{L}^*_{\alpha * X}(s)\} &= \sum_{n=0}^{\infty} \mathbb{P}[N(x) = n]\big(\mathbb{E}[-sW]\big)^n \\
&= \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{n!} \exp(-\lambda x)(M_W(-s))^n \\
&= \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{n!} \exp(-\lambda x)\Big(\frac{\alpha}{\alpha + s}\Big)^n \\
&= \exp(-\lambda x) \sum_{n=0}^{\infty} \frac{1}{n!}\Big(\frac{\lambda x \alpha}{\alpha + s}\Big)^n
\end{aligned}
$$

This simplification yields the setup the authors use in 2.1.1, with the exponential term independent of n taken out of the sum.

Proof of 2.1.2

From the above, we notice that

Finally, using the power series expansion of the exponential function,

$$
\exp(a) = \sum_{n=0}^{\infty} \frac{1}{n!} a^n
$$

The author's result follows:

$$
\begin{aligned}
\{\mathcal{L}^*_{\alpha * X}(s)\} &= \exp(-\lambda x) \exp\left(\frac{\lambda x \alpha}{\alpha + s}\right) \\
&= \exp\big(-\lambda x(1 - \frac{\alpha}{\alpha + s})\big) \\
&= \boxed{\exp\left(\frac{-\lambda x s}{\alpha + s}\right)}
\end{aligned}
$$

From this construction, the authors establish equation 2.2, the Laplace transform of a singular X_n conditioned on a fixed prior observation X_{n-1}. The below follows from the independence of ε_t from the cumulative process $\alpha * X_{n-1}$, yields

$$
\begin{aligned}
\mathbb{E}[\exp(-sX_n) \mid X_{n-1} = x] &= \mathbb{E}[\exp\big(-s(\alpha * X_{n-1} + \varepsilon_n)\big) \mid X_{n-1} = x] \\
&= \mathbb{E}[-s\varepsilon_t]\mathbb{E}[-s(\alpha * X_{n-1}) \mid X_{n-1} = x] \\
&= \boxed{\Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \exp\Big(\frac{-\lambda x s}{\alpha + s}\Big)} \quad \text{result 2.2}
\end{aligned}
$$

Then, the authors allow $\Phi_{X_n}(s)$ to be the Laplace transform of the PDF of X_n, i.e.

$$
\Phi_{X_n}(s) = \mathbb{E}[\exp(-sX_n)], \text{ for } s \geq 0
$$

From the above (conditional) LT, it follows by the Law of Total Expectation that

$$
\begin{aligned}
\mathbb{E}[\exp(-sX_n)] &= \mathbb{E}\Big[\mathbb{E}[\exp(-sX_n) \mid X_{n-1}]\Big] \\
\Phi_{X_n}(s) &= \mathbb{E}\Big[\Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \exp\Big(\frac{-\lambda X_{n-1} s}{\alpha + s}\Big)\Big] \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \mathbb{E}\Big[\exp\Big(-X_{n-1}\frac{\lambda s}{\alpha + s}\Big)\Big] \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \Phi_{X_{n-1}}\Big(\frac{\lambda s}{\alpha + s}\Big) \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \Phi_{X_{n-1}}\Big(\frac{\alpha p s}{\alpha + s}\Big), \text{ recalling } \lambda = \alpha p
\end{aligned}
$$

Setting up a simple recursive environment for $n = 2$

$$
\Phi_{X_2} = \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \Phi_{X_1}\Big(\frac{\alpha p s}{\alpha + s}\Big)
$$

Then, using the new Laplace variable $u = \alpha p s/\alpha + s$ (because we must substitute each time for consistency),

$$
\begin{aligned}
\Phi_{X_2} &= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \bigg(\Big(\frac{\alpha}{u+\alpha}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p u}{\alpha + u}\Big)\bigg) \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \bigg(\Big(\frac{\alpha}{\frac{\alpha p s}{\alpha+s} + \alpha}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p \frac{\alpha p s}{\alpha+s}}{\alpha + \frac{\alpha p s}{\alpha+s}}\Big)\bigg) \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \Big(\frac{\alpha}{\alpha\big[\frac{\alpha+s+ps}{\alpha+s}\big]}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p(\alpha p s)}{\alpha(\alpha+s) + \alpha p s}\Big) \\
&= \Big(\frac{\alpha}{s+\alpha}\Big)^{\nu} \Big(\frac{\alpha + s}{\alpha + p(1+s)}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^2 s}{\alpha + p(1+s)}\Big) \\
&= \Big(\frac{\alpha}{\alpha + p(1+s)}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^2 s}{\alpha + p(1+s)}\Big)
\end{aligned}
$$

From the above, the following general pattern emerges:

$$\Phi_{X_n}(s) = \Big(\frac{\alpha}{\alpha + s\sum_{k=0}^{n-1} p^k}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^n s}{\alpha + s\sum_{k=0}^{n-1} p^k}\Big)$$

Notice that if there is no serial correlation (i.e. $p = 0$) that the formula will be identical in each iterate (i.e. only dependent on α, s and $\Phi_{X_0}(\cdot)$).

Recall that in the above, $p \in [0, 1)$. Thus, we can employ the geometric sum equation:

$$\sum_{k=0}^{n-1} p^k = \frac{1 - p^n}{1 - p}$$

This allows us to simplify the previous into the form presented by the authors:

$$\begin{aligned}
\Phi_{X_n}(s) &= \Big(\frac{\alpha}{\alpha + s\frac{1-p^n}{1-p}}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^n s}{\alpha + s\frac{1-p^n}{1-p}}\Big) \\
&= \Big(\frac{\alpha}{\alpha + s\frac{1-p^n}{1-p}}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^n s}{\alpha + s\frac{1-p^n}{1-p}}\Big) \\
&= \Big(\frac{\alpha}{\frac{\alpha(1-p)+s(1-p^n)}{1-p}}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^n s}{\frac{\alpha(1-p)+s(1-p^n)}{1-p}}\Big) \\
&= \Big(\frac{\alpha(1-p)}{\alpha(1-p) + s(1-p^n)}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\alpha p^n (1-p) s}{\alpha(1-p) + s(1-p^n)}\Big) \\
&= \Big(\frac{\frac{1}{\theta}}{\frac{1}{\theta} + s(1-p^n)}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\frac{1}{\theta} p^n s}{\frac{1}{\theta} + s(1-p^n)}\Big), \text{ where } \theta = \frac{1}{\alpha(1-p)} \\
&= \Big(\frac{\frac{1}{\theta}}{\frac{1+\theta s(1-p^n)}{\theta}}\Big)^{\nu} \Phi_{X_0}\Big(\frac{\frac{1}{\theta} p^n s}{\frac{1+\theta s(1-p^n)}{\theta}}\Big) \\
&= \Big(\frac{1}{1 + \theta s(1-p^n)}\Big)^{\nu} \Phi_{X_0}\Big(\frac{p^n s}{1 + \theta s(1-p^n)}\Big) \\
&= \Big(\big(1 + \theta s(1-p^n)\big)^{-1}\Big)^{\nu} \Phi_{X_0}\Big(\frac{p^n s}{1 + \theta s(1-p^n)}\Big) \\
&= \boxed{\big(1 + \theta s(1-p^n)\big)^{-\nu} \Phi_{X_0}\Big(\frac{p^n s}{1 + \theta s(1-p^n)}\Big)} \quad (\star)
\end{aligned}$$

Yielding the cited result on Page 327. Directly, the above tends to $(1 + \theta s)^{-\nu}$ as $n \to \infty$, as the p^n terms decay to zero. From this, we can observe that:

$$(1 + \theta s)^{-\nu} = \Big(\frac{1}{1 + \theta s}\Big)^{\nu} = \Big(\frac{\theta^{-1}}{\theta^{-1} + s}\Big)^{\nu}$$

Now, recalling the Laplace transform of a Gamma distributed random variable with shape λ and rate ν (derived in the appendix and used earlier), it follows that X_n asymptotically approaches a Gamma(θ^{-1}, ν) distribution.

Indeed, in the parts to follow, the authors assume that the process $\{X_n\}_{n\in\mathbb{N}}$ is in equilibrium; hence, $\Phi_{X_n}(s) = (1+\theta s)^{-\nu}$ for all n.

Joint PDF : Proof of 2.3

Setup: 2.3.1

By using the Laplace Transform $\Phi_{X_n}(s)$ derived earlier, we can derive the LT of the joint PDF of X_n and X_{n-1} using the law of total expectation.

The authors use the suffix '2' in this case.

The following simplification follows from using Equation 2.2 and the nested/Markovian properties of $\{X_n\}$.

$$\begin{aligned}
\Phi_2(s_2, s_1) &= \mathbb{E}[\exp(-s_2 X_n - s_1 X_{n-1})] \\
&= \mathbb{E}\Big[\exp(-s_1 X_{n-1})\,\underbrace{\mathbb{E}[\exp(-s_2 X_n \mid X_{n-1})]}_{\text{Equation 2.2}}\Big] \\
&= \mathbb{E}\Big[\exp(-s_1 X_{n-1})\Big(\frac{\alpha}{s_2+\alpha}\Big)^{\nu}\exp\Big(-X_{n-1}\frac{\alpha p s_2}{\alpha+s_2}\Big)\Big] \\
&= \Big(\frac{\alpha}{s_2+\alpha}\Big)^{\nu}\mathbb{E}\Big[\exp\Big(-X_{n-1}\underbrace{(s_1+\frac{\alpha p s_2}{\alpha+s_2})}_{\text{Define as }\varrho}\Big)\Big] \\
&= \Big(\frac{\alpha}{s_2+\alpha}\Big)^{\nu}\mathbb{E}\Big[\exp\big(-X_{n-1}\varrho\big)\Big] \\
&= \Big(\frac{\alpha}{s_2+\alpha}\Big)^{\nu}\Phi_{X_{n-1}}(\varrho)
\end{aligned}$$

In the above, I set up the first line by definition of a bivariate Laplace transform. Then, in the second line, I expand the expectation using the Law of Total Expectation. This establishes a conditional of the form already established in Equation 2.2, which I substitute in the third line. Then, on the fourth line, I combine the exponential product into a single term and define a variable ϱ as the object of the LT argument $\{\mathcal{L}^*\}(\cdot)$. After substitution, in the fifth line, the system simplifies significantly and the final (sixth) line is in terms of the previously-defined LT Φ.

Then, the authors use the simplified equilibrium version of $\Phi_X(\cdot)$ to continue the simplification from this sixth line. The equilibrium version of the LT doesn't contain the $\Phi_{X_0}(\dots)$ term, and is given by

$$\Phi_{X_n}(s) = \big(1+\theta s\big)^{-\nu},\ \text{where } \theta = \frac{1}{\alpha(1-p)}$$

The above is assumed to be true for all n, since the previous simplification in ($\star$) tends to the above as $n \to \infty$ as discussed at the end of the previous section.

Proof of 2.3.2

Using the equilibrium version of $\Phi_{X_n}(s)$, we substitute

$$\Phi_2(s_2, s_1) = \Big(\frac{\alpha}{s_2 + \alpha}\Big)^{\nu}(1 + \theta\varrho)^{-\nu}$$

Now, we expand ϱ and express it in terms of θ rather than α.

$$\varrho = s_1 + \frac{\alpha p s_2}{\alpha + s_2} = s_1 + \frac{p s_2}{1 + \theta(1-p)s_2}$$

Now, we substitute this definition of ϱ and simplify using the fact that $1/\alpha = \theta(1-p)$

$$\begin{aligned}
\Phi_2(s_2, s_1) &= \Big(\frac{\alpha}{s_2 + \alpha}\Big)^{\nu}\Big(1 + \theta\Big(s_1 + \frac{p s_2}{1 + \theta(1-p)s_2}\Big)\Big)^{-\nu} \\
&= \Big(1 + \frac{s_2}{\alpha}\Big)^{-\nu}\Big(1 + \theta\Big(\frac{s_1 + \theta(1-p)s_1 s_2 + p s_2}{1 + \theta(1-p)s_2}\Big)\Big)^{-\nu} \\
&= \Big(1 + \theta(1-p)s_2\Big)^{-\nu}\Big(1 + \frac{\theta s_1 + \theta^2(1-p)s_1 s_2 + \theta p s_2}{1 + \theta(1-p)s_2}\Big)^{-\nu} \\
&= \Big((1 + \theta(1-p)s_2) \cdot \frac{1 + \theta s_2 - \theta p s_2 + \theta s_1 + \theta^2(1-p)s_1 s_2 + \theta p s_2}{1 + \theta(1-p)s_2}\Big)^{-\nu} \\
&= \big(1 + \theta(s_1 + s_2) + \theta^2(1-p)s_1 s_2\big)^{-\nu}
\end{aligned}$$

Yielding the cited result in (2.3).

Proof of 2.4

We now generalize this process for the LT of the joint pdf of $X_1, \dots, X_n$.

$$\Phi_n(s_n, s_{n-1}, \dots, s_1) = \mathbb{E}\big[\underbrace{\mathbb{E}[\exp(-s_n X_n) \mid X_{n-1}]}_{\text{Equation 2.2}}\exp(-s_{n-1}X_{n-1} - \cdots - s_1 X_1)\big]$$

Again, we identify the conditional component as Equation 2.2, derived earlier. Thus,

$$\Phi_n(s_n, s_{n-1}, \dots, s_1) = \mathbb{E}\Big[\big(1+\theta(1-p)s_n\big)^{-\nu}\exp\Big(-X_{n-1}\frac{\alpha p s_n}{\alpha + s_n}\Big)\exp(-s_{n-1}X_{n-1} - \cdots - s_1 X_1)\Big]$$

Now, following the simplification we did for ϱ earlier, it follows that

$$\begin{aligned}\Phi_n(s_n, s_{n-1}, \dots, s_1) &= \big(1+\theta(1-p)s_n\big)^{-\nu}\mathbb{E}\Big[\exp\Big(-X_{n-1}\frac{ps_n}{1+\theta(1-p)s_n}\Big)\exp(-s_{n-1}X_{n-1}-\dots-s_1X_1)\Big] \\ &= \big(1+\theta(1-p)s_n\big)^{-\nu}\mathbb{E}\Big[\exp(-X_{n-1}\Big(s_{n-1}+\frac{ps_n}{1+\theta(1-p)s_n}\Big)-s_{n-2}X_{n-2}-\dots-s_1X_1)\Big]\end{aligned}$$

Then, by definition of Laplace transform and our iterative structure (explored earlier in our lag-2 recursive example and in the derivation of $\Phi_2(s_2, s_1)$), we have that

$$\begin{aligned}\Phi_n(s_n, s_{n-1}, \dots, s_1) &= \big(1+\theta(1-p)s_n\big)^{-\nu}\mathbb{E}\Big[\exp\Big(-X_{n-1}\frac{ps_n}{1+\theta(1-p)s_n}\Big)\exp(-s_{n-1}X_{n-1}-\dots-s_1X_1)\Big] \\ &= \big(1+\theta(1-p)s_n\big)^{-\nu}\Phi_{n-1}\Big(s_{n-1}+\frac{ps_n}{1+\theta(1-p)s_n}, s_{n-2}, \dots, s_1\Big)\end{aligned}$$

Then, letting $G(s_j) = ps_j/1+\theta(1-p)s_j$, we arrive at the simplification in Step 2.4.

$$\Phi_n(s_n, s_{n-1}, \dots, s_1) = \big(1+\theta(1-p)s_n\big)^{-\nu}\Phi_{n-1}\Big(s_{n-1}+G(s_n), s_{n-2}, \dots, s_1\Big) \quad \square$$

Proof of 2.5: Primary Result of the Work

Notice that a recursive relationship exists in the above derivation. In this section, we will expand this to a general form for the LT of the joint PDF.

We now introduce a bit of nomenclature that the authors bring forward in the 'Introduction' section. Let $\mathbf{I}_n$ be the n-dimensional identity matrix, with 1s along the diagonal and zero elsewhere.

Further, let

$$\mathbf{S}_n = \begin{bmatrix} s_1 & s_2 & \dots & s_n \end{bmatrix} \times \mathbf{I}_n, \text{ and } \mathbf{V}_n = \begin{bmatrix} 1 & p^{1/2} & p & p^{3/2} & \dots & p^{|1-n|/2} \\ p^{1/2} & 1 & p^{1/2} & p & \dots & p^{|2-n|/2} \\ p & \ddots & \ddots & \ddots & \ddots & \vdots \\ p^{|n-1|/2} & \dots & \dots & \dots & \dots & 1 \end{bmatrix}, \text{ with } v_{ij} = p^{|i-j|/2}$$

The authors then define $\mathbf{A}_n = \mathbf{I}_n + \theta\mathbf{S}_n\mathbf{V}_n$ and claim that for all $n \geq 1$ that

$$\Phi_n(s_n, \dots, s_1) = \det(\mathbf{A}_n)^{-\nu} = \det(\mathbf{I}_n + \theta\mathbf{S}_n\mathbf{V}_n)^{-\nu}$$

Following the author's argument, I proceed by mathematical induction.

Base Case

In the base case, we have:

$$\det(\mathbf{A}_1) = (\begin{bmatrix}1\end{bmatrix} + \theta \begin{bmatrix}s_1\end{bmatrix} \begin{bmatrix}1\end{bmatrix})^{-\nu} = (1 + \theta s_1)^{-\nu}$$

Which matches our previous result.

'Inductive' Step

I tried for many hours to get this to work for general n… but I couldn't figure it out.

So instead let's look at $n = 2$ case (since we can compute 2×2 determinants by hand.)

To see the linear algebra in action, we consider the $n = 2$ case since the $n = 1$ case only includes scalars and trivially holds .

Here, we write $\mathbf{A}_2$ as

$$\begin{aligned}
\mathbf{A}_2 &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \theta \begin{bmatrix}s_1 & 0\\ 0 & s_2\end{bmatrix} \begin{bmatrix}1 & p^{1/2}\\ p^{1/2} & 1\end{bmatrix}\\
\mathbf{A}_2 &= \begin{bmatrix}1 + \theta s_1 & \theta s_1 p^{1/2}\\ \theta s_2 p^{1/2} & 1 + \theta s_2\end{bmatrix}
\end{aligned}$$

Then, the determinant to the power of $-\nu$ is computed as

$$\begin{aligned}
\det(\mathbf{A}_2)^{-\nu} &= \Big((1 + \theta s_1)(1 + \theta s_2) - (\theta s_2 p^{1/2} \theta s_1 p^{1/2})\Big)^{-\nu}\\
&= \left(1 + \theta s_1 + \theta s_2 + \theta^2 s_1 s_2 - \theta^2 s_1 s_2 p\right)^{-\nu}\\
&= \left(1 + \theta s_1 + \theta s_2 + \theta^2 (1 - p) s_1 s_2\right)^{-\nu}
\end{aligned}$$

This matches our previous derivation fo (2.3), and shows that the general solution will hold via induction.

Proof of 2.6

Now, the authors use the previously established theorem to write an equation for the LT of the joint density of X_n and X_{n+j} for lag j. Since the process is Markovian (and thus time-invariant), they utilize the LT of the joint PDF of 1 and $1 + j$.

Specifically, the authors claim that:

$$\Phi_{j+1}(s_{j+1}, 0, \ldots, 0, s_1) = \left(1 + \theta(s_1 + s_{j+1}) + \theta^2 (1 - p^j) s_1 s_{j+1}\right)^{-\nu}$$

This is an important result, as the un-transformed version forms the bivariate joint PDF for an observation X_n and a lagged observation X_{n+j}.

Proof

To prove this, we will use the result 2.5. We write $\mathbf{S}_{j+1} = \text{diag}(s_1, 0, \dots, 0, s_{j+1})$, and let $\mathbf{V}_{j+1}$ be as stated with $v_{ij} = p^{|i-j|/2}$.

Notice that $\mathbf{S}_{j+1}$ contains only two elements. We can thus write it as:

$$\mathbf{S}_n = s_1 e_1 e_1^\top + s_{j+1} e_{j+1} e_{j+1}^\top$$

And thus, when multiplied by $\mathbf{V}_{j+1}$, we can rearrange as follows:

$$\begin{aligned}
\mathbf{S}_n \mathbf{V}_n &= s_1 e_1 e_1^\top \mathbf{V}_{j+1} + s_{j+1} e_{j+1} e_{j+1}^\top \mathbf{V}_{j+1} \\
&= \begin{bmatrix} s_1 e_1 & s_{j+1} e_{j+1} \end{bmatrix} \begin{bmatrix} e_1^\top \mathbf{V}_{j+1} \\ e_{j+1}^\top \mathbf{V}_{j+1} \end{bmatrix} \\
&= \mathbf{U}\mathbf{W} \\
\text{Letting } \mathbf{W} = \begin{bmatrix} e_1^\top \mathbf{V}_{j+1} \\ e_{j+1}^\top \mathbf{V}_{j+1} \end{bmatrix}&, \mathbf{U} = \begin{bmatrix} s_1 e_1 & s_{j+1} e_{j+1} \end{bmatrix}
\end{aligned}$$

Notice that $\mathbf{U}_{(j+1)\times 2} \mathbf{W}_{2\times(j+1)}$ is a $(j+1) \times (j+1)$ matrix, hence it has a difficult to compute determinant. However, the Weinstein–Aronszajn identity (source) states that, for matrices $\mathbf{A}$ and $\mathbf{B}$ of dimensions $m \times n$ and $n \times m$ respectively, that:

$$\det(I_n + \mathbf{AB}) = \det(I_m + \mathbf{BA})$$

Recall that Equation (2.5) is of this exact structure. We can thus write

$$\Phi_{j+1}(s_{j+1}, 0, \dots, 0, s_1) = \det(\mathbf{I}_{j+1} + \theta \mathbf{S}_n \mathbf{V}_n)^{-\nu} = \det(\mathbf{I}_{j+1} + \theta \mathbf{U}\mathbf{W})^{-\nu} = \det(\mathbf{I}_2 + \theta \mathbf{W}\mathbf{U})^{-\nu}$$

The 2×2 computation of $\mathbf{WU}$ is rather direct.

$$\begin{aligned}
\mathbf{WU} &= \begin{bmatrix} e_1^\top \mathbf{V}_{j+1} \\ e_{j+1}^\top \mathbf{V}_{j+1} \end{bmatrix} \begin{bmatrix} s_1 e_1 & s_{j+1} e_{j+1} \end{bmatrix} \\
&= \begin{bmatrix} e_1^\top \mathbf{V}_{j+1} e_1 s_1 & e_1^\top \mathbf{V}_{j+1} e_{j+1} s_{j+1} \\ e_{j+1}^\top \mathbf{V}_{j+1} e_1 s_1 & e_{j+1}^\top \mathbf{V}_{j+1} e_{j+1} s_{j+1} \end{bmatrix} \\
&= \begin{bmatrix} v_{11} s_1 & v_{1(j+1)} s_{j+1} \\ v_{(j+1)1} s_1 & v_{(j+1)(j+1)} s_{j+1} \end{bmatrix} \\
&= \begin{bmatrix} p^{|1-1|/2} s_1 & p^{|1-(j+1)|/2} s_{j+1} \\ p^{|(j+1)-1|/2} s_1 & p^{|(j+1)-(j+1)|/2} s_{j+1} \end{bmatrix} \\
&= \begin{bmatrix} s_1 & p^{j/2} s_{j+1} \\ p^{j/2} s_1 & s_{j+1} \end{bmatrix}
\end{aligned}$$

From this simplification of $\mathbf{WU}$, we can compute the determinant

$$\begin{aligned}\Phi_{j+1}(s_{j+1}, 0, \dots, 0, s_1) &= \det(\mathbf{I}_2 + \theta\mathbf{WU})^{-\nu} \\ &= \det\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \theta\begin{bmatrix}s_1 & p^{j/2}s_{j+1}\\ p^{j/2}s_1 & s_{j+1}\end{bmatrix}\right)^{-\nu} \\ &= \det\left(\begin{bmatrix}1+\theta s_1 & \theta p^{j/2}s_{j+1}\\ \theta p^{j/2}s_1 & 1+\theta s_{j+1}\end{bmatrix}\right)^{-\nu} \\ &= \left((1+\theta s_1)(1+\theta s_{j+1}) - \theta^2 p^j s_1 s_{j+1}\right)^{-\nu} \\ &= \left(1+\theta(s_1+s_2) - \theta^2(1-p^j)s_1 s_{j+1}\right)^{-\nu} \;\; \square\end{aligned}$$

Yielding the cited result in 2.6

In the next document, we will use the reverse-Laplace transform of this conditional density function, alongside the previously-established marginal PDF of $\mathbf{X}_n$ to define the conditional density used in the second paper for model estimation.

Proof of 2.7: Conditional Density

We now use the fact that each X_n has $\Phi_{X_n}(s) = \left(1+\theta s\right)^{-\nu}$ (the so-called 'equilibrium version' used in 2.3.1) which defines a Gamma distribution with shape ν and scale θ^{-1} (equivalently, rate θ) with corresponding marginal probability density function:

$$f_{X_n}(x_n) = \frac{1}{\Gamma(\nu)\theta^\nu} x^{\nu-1} e^{-x/\theta}$$

As well as bivariate joint density (2.6.1)

$$f_{X_{n+j},X_n}(x,y) = \frac{1}{\Gamma(\nu)\theta^{\nu+1}(1-p^j)} \left(\frac{xy}{p^j}\right)^{(\nu-1)/2} \exp\left(\frac{-(x+y)}{\theta(1-p^j)}\right) I_{\nu-1}\left(\frac{2(p^j xy)^{1/2}}{\theta(1-p^j)}\right)$$

Consequently, the conditional density $f_{X_{t+j}|X_t}(x \mid y)$ can be computed as follows, letting the modified Bessel function component (rightmost term in the above) be written as C_{xy} for simplicity.

$$\begin{aligned}
f_{X_{t+j}|X_t}(x \mid y) &= \frac{f_{X_{t+j},X_t}(x,y)}{f_{X_t}(y)} \\
&= \frac{\frac{1}{\Gamma(\nu)\theta^{\nu+1}(1-p^j)}\Big(\frac{xy}{p^j}\Big)^{(\nu-1)/2}\exp\Big(\frac{-(x+y)}{\theta(1-p^j)}\Big)C_{xy}}{\frac{1}{\Gamma(\nu)\theta^{\nu}}y^{\nu-1}\exp\big(-\frac{y}{\theta}\big)} \\
&= \frac{\theta^{\nu-(\nu+1)}}{(1-p^j)}\Big(\frac{xy}{p^j}\Big)^{(\nu-1)/2}y^{-(\nu-1)}\exp\Big(\frac{-(x+y)}{\theta(1-p^j)}-\big(-\frac{y}{\theta}\big)\Big)C_{xy} \\
&= \frac{1}{\theta(1-p^j)}\Big(\frac{x}{p^jy}\Big)^{(\nu-1)/2}\exp\Big(-\frac{x+p^jy}{\theta(1-p^j)}\Big)I_{\nu-1}\Big(\frac{2(p^jxy)^{1/2}}{\theta(1-p^j)}\Big)
\end{aligned}$$

Recalling that $\theta = 1/(\alpha(1-p))$, we can write the above as:

$$f_{X_{t+j}|X_t}(x \mid y) = \frac{\alpha(1-p)}{(1-p^j)}\Big(\frac{x}{p^jy}\Big)^{(\nu-1)/2}\exp\Big(-\alpha(1-p)\frac{x+p^jy}{(1-p^j)}\Big)I_{\nu-1}\Big(\alpha(1-p)\frac{2(p^jxy)^{1/2}}{(1-p^j)}\Big)$$

Finally, letting $\zeta = \alpha(1-p)/(1-p^j)$, we have the derivation provided in the labeled (2.7) of the Second Paper.

$$\boxed{f_{X_{t+j}|X_t}(x \mid y) = \zeta\Big(\frac{x}{p^jy}\Big)^{(\nu-1)/2}\exp\big(-\zeta(x+p^jy)\big)I_{\nu-1}\big(2\zeta(p^jxy)^{1/2}\big)}$$

Where $I_r(x)$ is the modified Bessel function of the first kind and order r, given by (source):

$$I_r(x) = \sum_{m=0}^{\infty}\frac{1}{m!\,\Gamma(m+\alpha+1)}\Big(\frac{x}{2}\Big)^{2m+a}$$