Jiří Anděl; Georg Neuhaus

Extrapolation in fractional autoregressive models

Kybernetika, Vol. 34 (1998), No. 3, [309]--316

Persistent URL: http://dml.cz/dmlcz/135209

KYBERNETIKA — VOLUME 34 (1998), NUMBER 3, PAGES 309-316

EXTRAPOLATION IN FRACTIONAL AUTOREGRESSIVE MODELS

JIŘÍ ANDĚL AND GEORG NEUHAUS

The naïve and the least-squares extrapolation are investigated in the fractional autoregressive models of the first order. Some explicit formulas are derived for the one and two steps ahead extrapolation.

1. INTRODUCTION

Let $\{e_t, t \geq 1\}$ be i.i.d. random variables with a finite second moment (i.e., a strict white noise). Let λ be a function and X_0 a random variable independent of $\{e_1, e_2, \ldots\}$. Define

$$X_t = \lambda(X_{t-1}) + e_t, \quad t \geq 1. \tag{1.1}$$

If λ is a non-linear function then $\{X_t, t \geq 0\}$ is called a non-linear autoregressive process of the first order, briefly NLAR(1). Let $\gamma = Ee_t$. Assume that variables $X_0, \ldots, X_T$ are given and X_{T+m} for some $m \geq 1$ is to be extrapolated.

Define $X^*_{T|T} = X_T$ and

$$X^*_{T+m|T} = \lambda(X^*_{T+m-1|T}) + \gamma, \quad m \geq 1.$$

Then $X^*_{T+m|T}$ is called the naïve extrapolation of X_{T+m}. It is based on an analogy with the extrapolation in linear AR processes. It is useful to introduce functions

$$H_0(x) = x, \quad H_m(x) = \lambda[H_{m-1}(x)] + \gamma \ \text{ for } \ m \geq 1.$$

Then $X^*_{T+m|T} = H_m(X_T)$.

On the other hand, the least squares extrapolation $\widehat{X}_{T+m|T}$ of X_{T+m} is given by

$$\widehat{X}_{T+m|T} = E\{X_{T+m}|X_T, \ldots, X_0\} = E\{X_{T+m}|X_T\}$$

since the process $\{X_t\}$ is Markov. Assume that e_t has a density h. Introduce functions

$$K_0(x) = x, \quad K_m(x) = \int_{-\infty}^{\infty} K_{m-1}(y)\, h[y - \lambda(x)]\, \mathrm{d}y \ \text{ for } \ m \geq 1. \tag{1.2}$$

It is known that $\widehat{X}_{T+m|T} = K_m(X_T)$. A proof can be found in Tong [3], p. 346 for the case that $\{X_t\}$ is stationary. A modification of the proof without assumption of stationarity is straightforward.

It can be easily verified that $K_1(x) = H_1(x) = \lambda(x)+\gamma$. It means that $X^*_{T+1|T} = \widehat{X}_{T+1|T}$. However, $K_m(x) \neq H_m(x)$ generally holds if $m \geq 2$. The substantial difference between the naïve and the least squares extrapolations is that the naïve extrapolation depends on $\{e_t\}$ only through the expectation γ whereas the least squares extrapolation depends on the complete distribution of e_t.

Generally, it is very difficult to derive explicit formulas for $K_m(x)$ when $m \geq 2$. Such results are known only in very special cases (see Pemberton [2]). Usually, the calculation of least squares extrapolation is done only numerically.

2. FRACTIONAL AUTOREGRESSIVE MODELS

In the special case when

$$\lambda(x) = \frac{\sum_{j=0}^{p} a_j X_{t-1}^j}{\sum_{j=0}^{q} b_j X_{t-1}^j} \tag{2.1}$$

the model (1.1) is called the fractional autoregressive model of order 1, briefly FAR(1). In some special cases it was proposed by Jones [1] for non-linear extrapolation in meteorology (cf. Tong [3], p. 109 and p. 120). As for the choice (2.1), Tong [3] assumes that $0 \leq p \leq q+1 < \infty$, $a_p \neq 0$, $b_q \neq 0$. The function $\lambda(x)$ used in (2.1) can be extended to include also terms X_{t-s} for $s > 1$.

Jones [1] assumed that $e_t \sim N(0, \sigma^2)$ and investigated two choices of $\lambda(x)$, namely

$$\lambda(x) = \frac{x}{1+x^2} \quad \text{and} \quad \lambda(x) = \frac{x^3}{1+x^2}.$$

If e_t has a normal distribution then $\widehat{X}_{T+m|T}$ must be calculated numerically if $m \geq 2$. But explicit formulas for $\widehat{X}_{T+m|T}$ can be derived for example when e_t has density

$$h_r(x) = c_r(w^2 - x^2)^{r-1}, \quad -w < x < w, \quad r \geq 1 \tag{2.2}$$

where $w > 0$ is a parameter and

$$c_r = \frac{1}{2^{2r-1}w^{2r-1}B(r,r)}$$

is the normalizing constant. Since h_r is symmetric, $\gamma = Ee_t = 0$ and skewness of e_t is zero. The kurtosis $\alpha_4 = \mu_4/\mu_2^2 = 3(2r+1)/(2r+3) \to 3$ as $r \to \infty$.

In this paper we present some explicit formulas for $K_m(x)$ in FAR(1) models. The results were derived using the program package *Mathematica*.

3. THE FIRST MODEL

In this section we investigate the model (1.1) with $\lambda(x) = x/(1+x^2)$. This function is plotted in Figure 3.1. Let e_t have a density $h(x)$ such that $h(x) = h(-x)$ and $h(x_1) \geq h(x_2)$ for all $|x_1| \leq |x_2|$.

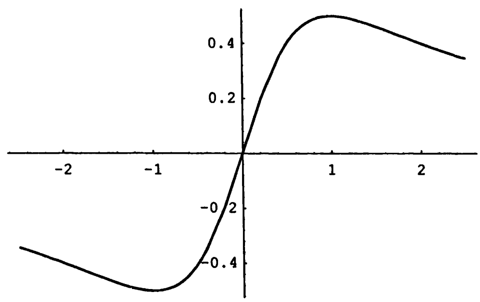

Fig. 3.1. Function $\lambda(x)$.

Theorem 3.1. For arbitrary $x \geq 0$ and $m \geq 0$ we have $K_m(-x) = -K_m(x)$ and $K_m(x) \geq 0$.

Proof. If $m = 0$ or $m = 1$ then the assertion is clear. Further we use complete induction. Let $m \geq 2$. Then

$$\begin{aligned} K_m(-x) &= \int_{-\infty}^{\infty} K_{m-1}(y)\, h[y - \lambda(-x)]\, \mathrm{d}y = -\int_{-\infty}^{\infty} K_{m-1}(-y)\, h[y + \lambda(x)]\, \mathrm{d}y \\ &= -\int_{-\infty}^{\infty} K_{m-1}(y)\, h[-y + \lambda(x)]\, \mathrm{d}y = -\int_{-\infty}^{\infty} K_{m-1}(y)\, h[y - \lambda(x)]\, \mathrm{d}y \\ &= -K_m(x). \end{aligned}$$

From Lemma A3 we obtain

$$K_m(x) = \int_{-\infty}^{\infty} K_{m-1}(y)\, h[y - \lambda(x)]\, \mathrm{d}y = \int_{-\infty}^{\infty} K_{m-1}[z + \lambda(x)]\, h(z)\, \mathrm{d}z \geq 0. \qquad \square$$

As a referee pointed out, the assertion of Theorem 3.1 can be generalized in the following way. Let the assumptions about e_t hold and let the function λ satisfy the condition $\lambda(-x) = \lambda(x)$. We show that this condition also suffices for antisymmetry of K_m. Denote by $L(x, \cdot)$ the conditional distribution of X_t given $X_{t-1} = x$. It follows from (1.1) that

$$L(x, A) = \int_A h[t - \lambda(x)]\, \mathrm{d}t.$$

The conditional distribution of X_{T+m} given $X_T = x$ is the m-times convolution L_m of the kernel L, i.e. $L_1 = L$ and for $m > 1$

$$L_m(x, A) = \int L_{m-1}(t, A)\, L(x, \mathrm{d}t).$$

From the formula for $L(x, A)$ one can see that the Lebesgue density $h_m[\cdot - \gamma(x)]$ of $L_m(x, \cdot)$ satisfies the relations $h_1 = h$ and for $m > 1$

$$h_m[t - \gamma(x)] = \int h_{m-1}[t - \gamma(y)]\, h[y - \gamma(x)]\, \mathrm{d}y.$$

Using the same arguments as in the proof of the first part in Theorem 3.1 it can be shown that

$$h_m[t - \gamma(-x)] = h_m[-t - \gamma(x)].$$

The antisymmetry $K_m(-x) = -K_m(x)$ then follows from the formula

$$K_m(x) = \int t h_m[t - \gamma(x)]\, \mathrm{d}t.$$

Theorem 3.2. For arbitrary $x \geq 0$ we have $K_2(x) \leq H_2(x)$.

Proof. We have

$$\begin{aligned} K_2(x) &= \int_0^\infty \{\lambda[\lambda(x) - z] + \lambda[\lambda(x) + z]\} h(z)\, \mathrm{d}z \\ &= 2\lambda(x) \int_0^\infty \frac{1 + \lambda^2(x) - z^2}{\{1 + [\lambda(x) - z]^2\}\{1 + [\lambda(x) + z]^2\}} h(z)\, \mathrm{d}z. \end{aligned}$$

Using Lemma A1 we get

$$K_2(x) \leq 2\lambda(x) \int_0^\infty \frac{1}{1 + \lambda^2(x)} h(z)\, \mathrm{d}z = \lambda[\lambda(x)] = H_2(x). \qquad \square$$

We have

$$H_2(x) = \frac{x(1 + x^2)}{1 + 3x^2 + x^4}.$$

The function $K_2(x)$ depends on $h(x)$. Write $K_{2,r}(x)$ instead of $K_2(x)$ when $h(x) = h_r(x)$. Further, write λ instead of $\lambda(x)$. Then

$$K_{2,1}(x) = \frac{1}{4w} \ln \frac{1 + (\lambda + w)^2}{1 + (\lambda - w)^2}.$$

For simplicity, consider the case $w = 2$. Then we get

$$\begin{aligned} K_{2,2}(x) &= \frac{0.046\,875}{(1 + x^2)^2} \Big\{ 8x(1 + x^2) - 4x(1 + x^2)[\operatorname{arctg}(2 - \lambda) + \operatorname{arctg}(2 + \lambda)] \\ &\quad + (5 + 9x^2 + 5x^4) \ln \frac{1 + (\lambda + 2)^2}{1 + (\lambda - 2)^2} \Big\}, \end{aligned}$$

$$\begin{aligned} K_{2,3}(x) &= \frac{0.004\,882\,81}{(1 + x^2)^4} \Big\{ 8x(29 + 84x^2 + 84x^4 + 29x^6) \\ &\quad - 24x(5 + 14x^2 + 14x^4 + 5x^6)[\operatorname{arctg}(2 - \lambda) + \operatorname{arctg}(2 + \lambda)] \\ &\quad + 3(25 + 86x^2 + 123x^4 + 86x^6 + 25x^8) \ln \frac{1 + (\lambda + 2)^2}{1 + (\lambda - 2)^2} \Big\}. \end{aligned}$$

These functions are plotted in Figure 3.2. In the legend the function $K_{2,r}(x)$ is denoted as $K2r(x)$. Figure 3.2 confirms that $H_2(x) \geq K_{2,r}(x)$ for $x \geq 0$ as proved in Theorem 3.2.

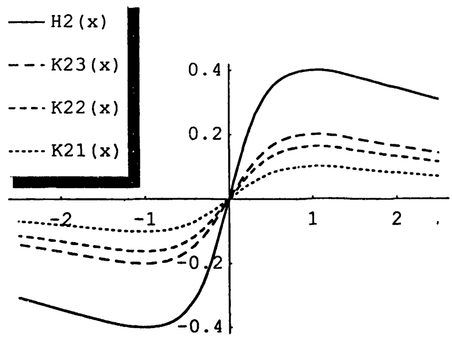

Fig. 3.2. Functions $H_2(x)$ and $K_{2,r}(x)$.

4. THE SECOND MODEL

Here we consider again the model (1.1) but this time with $\lambda(x) = x^3/(1+x^2)$ (see Figure 4.1). Assume also here that $h(x) = h(-x)$ for all x and $h(x_1) \geq h(x_2)$ for all $|x_1| \leq |x_2|$.

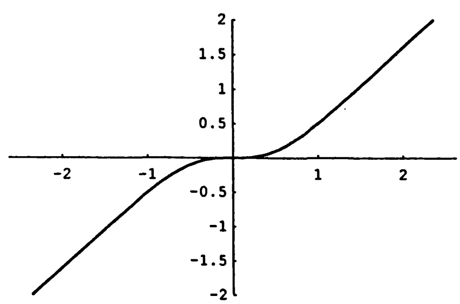

Fig. 4.1. Function $\lambda(x)$.

Theorem 4.1. For arbitrary $x \geq 0$ and $m \geq 0$ we have $K_m(-x) = -K_m(x)$ and $K_m(x) \geq 0$.

Proof. The assertion can be proved in the same way as Theorem 3.1. It follows also from the remark to the proof of Theorem 3.1. □

Theorem 4.2. For arbitrary $x \geq 0$ we have $K_2(x) \geq H_2(x)$.

Proof. Similarly as in the proof of Theorem 3.2 we get

$$K_2(x) = 2\int_0^\infty \frac{\lambda^3(x) + 3\lambda(x)z^2 + \lambda(x)[\lambda^2(x) - z^2]^2}{\{1 + [\lambda(x) - z]^2\}\{1 + [\lambda(x) + z]^2\}} \, dz.$$

Lemma A2 implies that

$$K_2(x) \leq 2\int_0^\infty \frac{\lambda^3(x) + \lambda^5(x)}{[1 + \lambda^2(x)]^2} \, dz = \frac{\lambda^3(x)}{1 + \lambda^2(x)} = \lambda[\lambda(x)] = H_2(x). \qquad \square$$

In this case

$$H_2(x) = \frac{x^9}{1 + 3x^2 + 3x^4 + 2x^6 + x^8}.$$

Write again λ instead of $\lambda(x)$. Choosing $h(x) = h_r(x)$ and $w = 2$ we obtained

$$\begin{aligned}
K_{2,1}(x) &= \lambda + \frac{1}{8}\ln\frac{1 + (\lambda - 2)^2}{1 + (\lambda + 2)^2}, \\
K_{2,2}(x) &= \frac{0.015\,625}{(1 + x^2)^2}\Big\{40x^3(1 + x^2) + 12x^3(1 + x^2)[\operatorname{arctg}(2 - \lambda) + \operatorname{arctg}(2 + \lambda)] \\
&\quad + 3(5 + 10x^2 + 5x^4 - x^6)\ln\frac{1 + (\lambda - 2)^2}{1 + (\lambda + 2)^2}\Big\}, \\
K_{2,3}(x) &= \frac{0.000\,976\,562}{(1 + x^2)^4}\Big\{8x^3(-17 - 51x^2 - 51x^4 - 2x^6 + 15x^8) \\
&\quad + 120x^3(5 + 15x^2 + 15x^4 + 4x^6 - x^8)[\operatorname{arctg}(2 - \lambda) + \operatorname{arctg}(2 + \lambda)] \\
&\quad + 15(25 + 100x^2 + 150x^4 + 86x^6 - 3x^8 - 14x^{10} + x^{12})\ln\frac{1 + (\lambda - 2)^2}{1 + (\lambda + 2)^2}\Big\}.
\end{aligned}$$

The functions $H_2(x)$, $K_{2,1}(x)$, $K_{2,2}(x)$, and $K_{2,3}(x)$ are plotted in Figure 4.2.

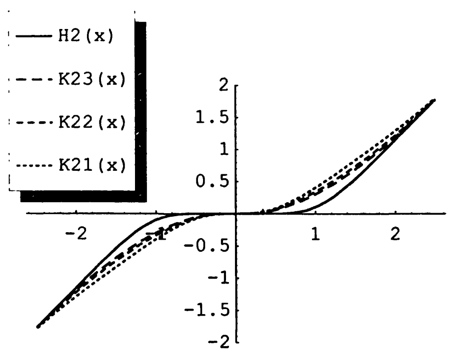

Fig. 4.2. Functions $H_2(x)$ and $K_{2,r}(x)$.

APPENDIX

Lemma A1. Let $a \in [-0.5, 0.5]$. Define

$$g(x) = \frac{1 + a^2 - x^2}{[1 + (a - x)^2]\,[1 + (a + x)^2]}.$$

Then $g(x) < g(0)$ for all $x \neq 0$.

Proof. The function g is symmetric. A straightforward calculation gives that $g'(x) = 0$ only for $x = x_1 = 0$ and for $x = x_{23} = \pm\sqrt{1 + a^2 + 2\sqrt{1 + a^2}}$. Since $g''(0) = 2(-3 - 8a^2 - 6a^4 + a^8)/(1 + a^2)^6 < 0$, g has maximum at $x = 0$. Similarly, g has minimum at x_2 and x_3. We have $g(0) > 0$, $g(x_2) = g(x_3) < 0$, $\lim g(x) = 0$ as $x \to \pm\infty$. Thus $g(0)$ is the global maximum. □

Lemma A2. Let $a \in (0, 0.5]$. Define

$$u(x) = \frac{a^3 + 3ax^2 + a(a^2 - x^2)^2}{[1 + (a - x)^2]\,[1 + (a + x)^2]}.$$

Then $u(0) < u(x)$ for all $x \neq 0$.

Proof. We have $u'(x) = 0$ only for $x = x_1 = 0$ and for $x = x_{23} = \pm\sqrt{1 + a^2 + 2\sqrt{1 + a^2}}$. Further, $u''(0) > 0$, $u''(x_2) < 0$, $u''(x_3) < 0$, and $\lim u(x) = a$ as $x \to \pm\infty$. Since $u(0) = a^3/(1 + a^2) < a$, the global minimum of $u(x)$ is $u(0)$. □

Lemma A3. Let $a > 0$. Let L be a function such that $L(x) \geq 0$ and $L(-x) = -L(x)$ for all $x > 0$. Let h be a density such that $h(x) = h(-x)$ for all x and $h(x_1) \geq h(x_2)$ for all $|x_1| \leq |x_2|$. Then

$$\int_{-\infty}^{\infty} L(x + a)\, h(x)\, \mathrm{d}x \geq 0.$$

Proof. We have

$$\begin{aligned}
\int_{-\infty}^{\infty} L(x + a)\, h(x)\, \mathrm{d}x &= \int_{-\infty}^{\infty} L(t)\, h(t - a)\, \mathrm{d}t \\
&= \int_{-\infty}^{0} L(t)\, h(t - a)\, \mathrm{d}t + \int_{0}^{\infty} L(t)\, h(t - a)\, \mathrm{d}t \\
&= \int_{0}^{\infty} L(t)[h(t - a) - h(-t - a)]\, \mathrm{d}t.
\end{aligned}$$

The last integral is non-negative since for $t \geq 0$ we have $L(t) \geq 0$ and $|t - a| \leq |-t - a| = t + a$. □

ACKNOWLEDGEMENTS

The research of the first author was supported by grant 201/97/1176 from the Grant Agency of the Czech Republic. The authors are grateful to an anonymous referee for suggesting the generalization of Theorem 3.1.

(Received September 28, 1997.)

REFERENCES

[1] R. H. Jones: An experiment in non-linear prediction. J. Appl. Meteorol. *4* (1965), 701–705.
[2] J. Pemberton: Exact least squares multi-step prediction from nonlinear autoregressive models. J. Time Ser. Anal. *8* (1987), 443–448.
[3] H. Tong: Non-linear Time Series. Clarendon Press, Oxford 1990.
[4] Wolfram Research, Inc.: Mathematica, Version 2.2, Champaign, Illinois 1994.

Prof. RNDr. Jiří Anděl, DrSc., Charles University, Faculty of Mathematics and Physics, Sokolovská 83, 186 00 Praha 8. Czech Republic.
e-mail: andel@karlin.mff.cuni.cz

Prof. Dr. Georg Neuhaus, University of Hamburg, Department of Mathematical Stochastics, Bundesstr. 55, 2000 Hamburg. Federal Republic of Germany.
e-mail: neuhaus@math.uni-hamburg.de