Kybernetika

Jiří Anděl; Václav Dupač

Extrapolations in non-linear autoregressive processes

Kybernetika, Vol. 35 (1999), No. 3, [383]--389

Persistent URL: http://dml.cz/dmlcz/135294

KYBERNETIKA — VOLUME 35 (1999), NUMBER 3, PAGES 383–389

EXTRAPOLATIONS IN NON-LINEAR AUTOREGRESSIVE PROCESSES

JIŘÍ ANDĚL AND VÁCLAV DUPAČ

We derive a formula for m-step least-squares extrapolation in non-linear AR(p) processes and compare it with the naïve extrapolation. The least-squares extrapolation depends on the distribution of white noise. Some bounds for it are derived that depend only on the expectation of white noise. An example shows that in general case the difference between both types of extrapolation can be very large. Further, a formula for least-squares extrapolation in multidimensional non-linear AR(p) process is derived.

1. INTRODUCTION

Let $e_1, e_2, \ldots$ be i.i.d. random variables with finite second moment. Assume that e_t has a density h. Let $X_0, X_{-1}, \ldots, X_{-p+1}$ be random variables independent of $\{e_t, t \geq 1\}$. Define $\gamma = Ee_t$ and

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

where λ is a Borel measurable (generally a non-linear) function such that $EX_t^2 < < \infty$. Then $\{X_t, t \geq 1\}$ is called non-linear autoregressive process of order p, briefly NLAR(p).

Assume that for a $t \geq 1$ the variables $\{X_{t-s}, s \geq 0\}$ are given and X_{t+m} is to be extrapolated. The naïve extrapolation $X^*_{t+m|t}$ is defined as follows. Let $X^*_{t+m|t} = X_{t+m}$ for $m \leq 0$ and

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

Denote $z = (z_1, \ldots, z_p)'$ and introduce functions

$$H_m(z) = \begin{cases} z_{-m+1} & \text{for} \quad m = 0, -1, \ldots, -p+1, \\ \lambda[H_{m-1}(z), \ldots, H_{m-p}(z)] + \gamma & \text{for} \quad m \geq 1. \end{cases} \tag{1.2}$$

Then

$$X^*_{t+m|t} = H_m(X_t, X_{t-1}, \ldots, X_{t-p+1}).$$

Note that $H_m(z) = H_1[H_{m-1}(z), \ldots, H_{m-p}(z)]$.

The least-squares (LS) extrapolation $\hat{X}_{t+m|t}$ of the variable X_{t+m} given $\{X_{t-s}, s \geq 0\}$ is defined by

$$\hat{X}_{t+m|t} = E(X_{t+m}|X_t, X_{t-1}, \dots, X_{-p+1}).$$

Since the model (1.1) is Markovian of the pth order, we see that

$$\hat{X}_{t+m|t} = E(X_{t+m}|X_t, X_{t-1}, \dots X_{t-p+1}).$$

Then the variable $\hat{X}_{t+m|t}$ can be written in the form

$$\hat{X}_{t+m|t} = K_{m,t}(X_t, X_{t-1}, \dots X_{t-p+1})$$

where $K_{m,t}$ is a function.

2. PROPERTIES OF EXTRAPOLATIONS

Calculation of LS extrapolation can be based on the following theorem.

Theorem 2.1. Functions $K_{m,t}$, $m \geq 0$, $t \geq 0$, are independent of t. They satisfy, with subscript t already dropped, the relation

$$K_m(z_1, \dots, z_p) = \int K_{m-1}(w, z_1, \dots, z_{p-1})\, h[w - \lambda(z_1, \dots, z_p)]\, \mathrm{d}w, \quad m \geq 1. \tag{2.1}$$

Proof. Introduce p-vectors $\boldsymbol{X}_t = (X_t, \dots, X_{t-p+1})'$, $t \geq 0$, $\boldsymbol{u} = (1, 0, \dots, 0)'$ and a transformation

$$T(\boldsymbol{z}) = (\lambda(\boldsymbol{z}), z_1, z_2, \dots, z_{p-1})', \quad \boldsymbol{z} \in \mathbb{R}^p.$$

We have

$$\boldsymbol{X}_{t+1} = T(\boldsymbol{X}_t) + e_{t+1}\boldsymbol{u}, \quad t \geq 0;$$

hence, $\{\boldsymbol{X}_t\}$ is a discrete parameter p-dimensional Markov process with the initial distribution of $\boldsymbol{X}_0$ and with a stationary transition distribution function

$$F(\boldsymbol{z}, \boldsymbol{y}) = H(y_1 - \lambda(\boldsymbol{z}))\, I(y_2 - z_1) \dots I(y_p - z_{p-1})$$

where H is the distribution function of e_1 and

$$I(y) = \begin{cases} 1 & \text{for} \quad y \geq 0, \\ 0 & \text{for} \quad y < 0. \end{cases}$$

Denote by $F^{(m)}(\boldsymbol{z}, \boldsymbol{y})$ the m-step transition distribution function of the process. As $K_{m,t}(\boldsymbol{z}) = E(\boldsymbol{X}'_{t+m}\boldsymbol{u}|\boldsymbol{X}_t = \boldsymbol{z})$, the independence of t follows. Further,

$$K_m(\boldsymbol{z}) = \int_{\mathbb{R}^p} \boldsymbol{y}'\boldsymbol{u} F^{(m)}(\boldsymbol{z}, \mathrm{d}\boldsymbol{y}) = \int_{\mathbb{R}^p} \int_{\mathbb{R}^p} \boldsymbol{y}'\boldsymbol{u} F(\boldsymbol{z}, \mathrm{d}\boldsymbol{w}) F^{(m-1)}(\boldsymbol{w}, \mathrm{d}\boldsymbol{y})$$

$$= \int_{\mathbb{R}^p} K_{m-1}(\boldsymbol{w}) F(\boldsymbol{z}, \mathrm{d}\boldsymbol{w}) = \int_{-\infty}^{\infty} K_{m-1}(w_1, z_1, z_2, \dots, z_{p-1}) h(w_1 - \lambda(\boldsymbol{z}))\, \mathrm{d}w_1,$$

proving thus (2). □

Theorem 1 for $p = 1$ and with an unnecessary stationarity assumption of the process $\{X_t\}$ is given in Tong [3].

It is clear that $K_0(\boldsymbol{z}) = z_1$. Moreover, it follows immediately from our definitions that $H_1(\boldsymbol{z}) = K_1(\boldsymbol{z}) = \lambda(\boldsymbol{z}) + \gamma$. If $m \geq 2$ then $K_m \neq H_m$ generally holds, since

$$\begin{aligned} H_2(\boldsymbol{z}) &= \lambda[\lambda(\boldsymbol{z}) + \gamma, z_1, \ldots, z_{p-1}] + \gamma, \\ K_2(\boldsymbol{z}) &= \int \lambda[y + \lambda(\boldsymbol{z}), z_1, \ldots, z_{p-1}]\, h(y)\, \mathrm{d}y + \gamma. \end{aligned}$$

In special cases, $K_m = H_m$ holds even for some $m \geq 2$. A typical situation is described in the following theorem.

Theorem 2.2. Let $p \geq 2$ and let

$$X_t = bX_{t-1} + \varphi(X_{t-2}, \ldots, X_{t-p}) + e_t \tag{2.2}$$

where φ is a Borel measurable function. Then $K_2(\boldsymbol{z}) = H_2(\boldsymbol{z})$ holds for all $\boldsymbol{z}$.

Proof. In our model (2.2) we have

$$K_1(\boldsymbol{z}) = H_1(\boldsymbol{z}) = bz_1 + \varphi(z_2, \ldots, z_p) + \gamma.$$

Using (2.1) we get

$$\begin{aligned} K_2(\boldsymbol{z}) &= \int [by + \varphi(z_1, \ldots, z_{p-1}) + \gamma] h[y - bz_1 - \varphi(z_2, \ldots, z_p)]\, \mathrm{d}y \\ &= \int [bu + b^2 z_1 + b\varphi(z_2, \ldots, z_p) + \varphi(z_1, \ldots, z_{p-1}) + \gamma]\, h(u)\, \mathrm{d}u \\ &= b\gamma + b^2 z_1 + b\varphi(z_2, \ldots, z_p) + \varphi(z_1, \ldots, z_{p-1}) + \gamma. \end{aligned}$$

Since

$$X_{t+2} = bX_{t+1} + \varphi(X_t, \ldots, X_{t-p+2}) + e_{t+2}$$

and the naïve extrapolation of X_{t+1} is

$$X^*_{t+1|t} = bX_t + \varphi(X_{t-1}, \ldots, X_{t-p+1}) + \gamma$$

we get

$$\begin{aligned} X^*_{t+2|t} &= bX^*_{t+1|t} + \varphi(X_t, \ldots, X_{t-p+1}) + \gamma \\ &= b\gamma + b^2 X_t + b\varphi(X_{t-1}, \ldots, X_{t-p+1}) + \varphi(X_t, \ldots, X_{t-p+1}) + \gamma. \end{aligned}$$

From this formula for $X^*_{t+2|t}$ it is clear that $K_2(\boldsymbol{z}) = H_2(\boldsymbol{z})$. □

Consider again the general model (1.1). It is clear that the naïve extrapolation depends on the white noise $\{e_t\}$ only through its expectation γ whereas the LS extrapolation depends on the complete distribution of e_t. It can be of some use to have some bounds for LS extrapolation such that they depend also only on γ. We derive such results for a class of functions $\lambda(\boldsymbol{z})$. First of all, we introduce auxiliary assertions.

Lemma 2.3. Assume that $f(\boldsymbol{z})$ is a concave function such that it is non-decreasing in each variable. Let $g_1(\boldsymbol{z}), \ldots, g_p(\boldsymbol{z})$ be concave functions. Then the function $f[g_1(\boldsymbol{z}), \ldots, g_p(\boldsymbol{z})]$ is concave.

Proof. The assertion can be proved directly. □

Lemma 2.4. Let ξ be a random variable such that $a \leq \xi \leq b$ where $-\infty < a < b < \infty$. Define $p = (E\xi - a)/(b - a)$, $q = 1 - p$. If ϕ is a convex function on $[a, b]$ then

$$E\phi(\xi) \leq q\phi(a) + p\phi(b). \tag{2.3}$$

Proof. See Kall and Wallace [2], p. 168. □

The inequality (2.3) is called Edmundson–Madansky upper bound.

Lemma 2.5. Let $\lambda(\boldsymbol{z})$ be a concave function. Then the functions $H_m(\boldsymbol{z})$, $m \geq -p + 1$, are concave and non-decreasing in each variable.

Proof. Theorem can be proved by complete induction using Lemma 2.3. □

Lemma 2.6. Let

$$\begin{aligned} U_1(\boldsymbol{z}) &= H_1(\boldsymbol{z}), \\ U_m(\boldsymbol{z}) &= U_{m-1}[\gamma + \lambda(\boldsymbol{z}), z_1, \ldots, z_{p-1}], \quad m \geq 2. \end{aligned}$$

If $\lambda(\boldsymbol{z})$ is a concave function non-decreasing in each variable then the functions $U_m(\boldsymbol{z})$, $m \geq 1$, are concave.

Proof. The assertion follows from Lemma 2.3 using complete induction. □

By the way, it is easy to see that $U_2(\boldsymbol{z}) = H_2(\boldsymbol{z})$.

Theorem 2.7. Let $\lambda(\boldsymbol{z})$ be a concave function non-decreasing in each variable. Then $K_m(\boldsymbol{z}) \leq U_m(\boldsymbol{z})$, $m \geq 1$.

Proof. We use complete induction. We know already that $K_1(\boldsymbol{z}) = H_1(\boldsymbol{z}) = U_1(\boldsymbol{z})$. Let $m = 2$. Our assumptions ensure that the function $r(y) = \lambda[y + \lambda(\boldsymbol{z}), z_1, \ldots, z_{p-1}]$ is concave for arbitrary fixed $\boldsymbol{z}$. Jensen inequality gives

$$\begin{aligned} K_2(\boldsymbol{z}) &= \gamma + \int \lambda[y + \lambda(\boldsymbol{z}), z_1, \ldots, z_{p-1}]\, h(y)\,\mathrm{d}y \\ &\leq \gamma + \lambda[\gamma + \lambda(\boldsymbol{z}), z_1, \ldots, z_{p-1}] \\ &= H_2(\boldsymbol{z}) = U_2(\boldsymbol{z}). \end{aligned}$$

For $m > 2$ it follows from the induction assumption, Lemma 2.6 and Jensen inequality that

$$\begin{aligned} K_m(\boldsymbol{z}) &= \int K_{m-1}[y+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}]\, h(y)\, \mathrm{d}y \\ &\leq \int U_{m-1}[y+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}]\, h(y)\, \mathrm{d}y \\ &\leq U_{m-1}[\gamma+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}] = U_m(\boldsymbol{z}). \end{aligned}$$ □

Theorem 2.8. Let $a \leq e_t \leq b$ where $-\infty < a < b < \infty$. Define $p = (\gamma - a)/(b-a)$, $q = 1-p$ and

$$\begin{aligned} J_1(\boldsymbol{z}) &= \lambda(\boldsymbol{z}), \\ J_m(\boldsymbol{z}) &= qJ_{m-1}[a+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}] + pJ_{m-1}[a+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}], \quad m \geq 2, \\ L_m(\boldsymbol{z}) &= \gamma + J_m(\boldsymbol{z}), \quad m \geq 1. \end{aligned}$$

If $\lambda(\boldsymbol{z})$ is a concave function non-decreasing in each variable then J_m, $m \geq 1$, are concave functions and $L_m(\boldsymbol{z}) \leq K_m(\boldsymbol{z})$ for $m \geq 2$.

Proof. It is clear that $J_m(\boldsymbol{z})$ are concave functions. From Lemma 2.4 we have

$$K_2(\boldsymbol{z}) = \gamma + \int_a^b \lambda[y+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}]\, h(y)\, \mathrm{d}y \geq \gamma + J_2(\boldsymbol{z}) = L_2(\boldsymbol{z}).$$

Similarly, by complete induction we obtain for $m \geq 3$ that

$$\begin{aligned} K_m(\boldsymbol{z}) &= \int_a^b K_{m-1}[y+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}] h(y)\, \mathrm{d}y \\ &\geq \gamma + \int_a^b J_{m-1}[y+\lambda(\boldsymbol{z}), z_1, \dots, z_{p-1}] h(y)\, \mathrm{d}y \\ &\geq \gamma + J_m(\boldsymbol{z}) = L_m(\boldsymbol{z}). \end{aligned}$$ □

3. AN EXAMPLE

Let $v > 0$. Assume that $e_t \sim R(0, v)$ and define

$$X_t = \sqrt{X_{t-1}X_{t-2}} + e_t, \quad t \geq 1$$

where X_0 and X_1 are non-negative random variables. Then $\gamma = v/2$. From $X_{t+1} = \sqrt{X_t X_{t-1}} + e_{t+1}$, $X_{t+2} = \sqrt{X_{t+1}X_t} + e_{t+2}$ we get

$$H_1(z_1, z_2) = \sqrt{z_1 z_2} + \gamma, \quad H_2(z_1, z_2) = \sqrt{(\sqrt{z_1 z_2} + \gamma)\, z_1} + \gamma.$$

It was already mentioned that $K_1(z_1, z_2) = H_1(z_1, z_2)$. From (2.1) we obtain

$$K_2(z_1, z_2) = \gamma + \frac{2}{3v}\sqrt{z_1}\left[(v + \sqrt{z_1 z_2})^{3/2} - (z_1 z_2)^{3/4}\right].$$

For $v = 10$ the functions H_1, H_2, K_2, and $D_2 = K_2 - H_2$ are plotted in Figures 1, 2, 3, and 4, respectively.

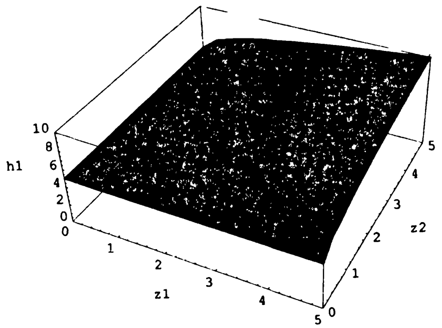

Fig. 1.

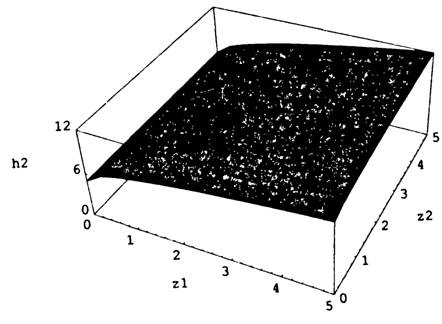

Fig. 2.

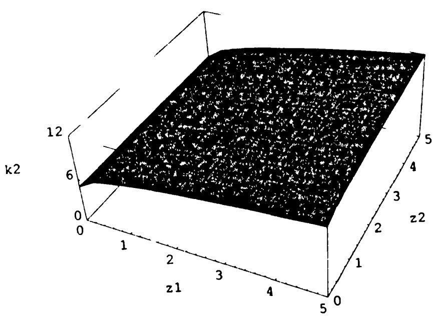

Fig. 3.

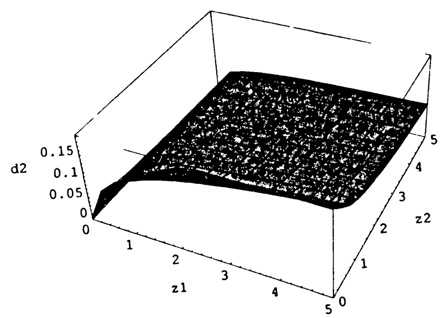

Fig. 4.

Consider the difference

$$\delta(v) = K_2(1,1) - H_2(1,1) = \sqrt{1 + \frac{v}{2}} - \frac{2}{3v}\left[(v+1)^{3/2} - 1\right].$$

It is clear that $\delta(v) \to \infty$ as $v \to \infty$. This is an elementary example that the difference between the least-squares and the naïve extrapolations can be arbitrary large. A similar example for an NLAR(1) process is introduced in Anděl [1].

4. MULTIDIMENSIONAL NLAR(p) PROCESSES

Now, let $e_1, e_2, \ldots$ be i.i.d. random q-vectors with finite second-order moments. Assume that e_t has a density h. Define $\gamma = Ee_t$. Let $X_0, X_{-1}, \ldots, X_{-p+1}$ be random q-vectors, independent of $\{e_t, t \geq 1\}$. Let λ be a Borel measurable function from $\mathbb{R}^{pq}$ into $\mathbb{R}^q$. For q-vectors $z_1, \ldots, z_p, \lambda(z_1, \ldots, z_p)$ or $\lambda(\mathcal{Z})$ will denote the function value of λ at point $\mathcal{Z} = vec(z_1, \ldots, z_p)$. Define

$$X_t = \lambda(X_{t-1}, \ldots, X_{t-p}) + e_t, \quad t \geq 1,$$

and assume that the second order moments of $\boldsymbol{X}_t$ are finite for all $t \geq 1$. Thus $\{\boldsymbol{X}_t, t \geq 1\}$ is a q-dimensional NLAR(p) process.

Introduce functions $H_m : \mathbb{R}^{pq} \to \mathbb{R}^q$ by

$$H_m(\mathcal{Z}) = \begin{cases} z_{-m+1} & \text{for} \quad m = 0, -1, \ldots, -p+1, \\ \lambda[H_{m-1}(\mathcal{Z}), \ldots, H_{m-p}(\mathcal{Z})] + \gamma & \text{for} \quad m \geq 1. \end{cases}$$

Then $\boldsymbol{X}^*_{t+m|t} = H_m(\boldsymbol{X}_t, \ldots, \boldsymbol{X}_{t-p+1})$ is the naïve extrapolation of $\boldsymbol{X}_{t+m}$ given $\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots$.

Theorem 4.1. Define functions $K_m : \mathbb{R}^{pq} \to \mathbb{R}^q$ by

$$\begin{aligned} K_0(z_1, \ldots, z_p) &= z_1, \\ K_m(z_1, \ldots, z_p) &= \int K_{m-1}(w, z_1, \ldots, z_{p-1})\, h[w - \lambda(z_1, \ldots, z_p)]\, \mathrm{d}w, \quad m \geq 1. \end{aligned}$$

Then

$$\hat{\boldsymbol{X}}_{t+m|t} = K_m(\boldsymbol{X}_t, \ldots, \boldsymbol{X}_{t-p+1})$$

is the least-squares extrapolation of $\boldsymbol{X}_{t+m}$ given $\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots$.

Proof. Introduce $\mathcal{X}_t = vec(\boldsymbol{X}_t, \ldots, \boldsymbol{X}_{t-p+1})$ and

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{I} \\ \cdots \\ \boldsymbol{0} \end{pmatrix},$$

where $\boldsymbol{I}$ is the $q \times q$ identity matrix and $\boldsymbol{0}$ the $(p-1)q \times q$ zero matrix. With $T(\mathcal{Z}) = vec[\lambda(\mathcal{Z}), z_1, \ldots, z_{p-1}]$, $\mathcal{Z} \in \mathbb{R}^{pq}$, we have $\mathcal{X}_{t+1} = T(\mathcal{X}_t) + \boldsymbol{u}e_{t+1}$. The rest of the proof is quite the same as that of Theorem 2.1. □

ACKNOWLEDGEMENT

The research was supported by grant 201/97/1176 from the Grant Agency of the Czech Republic.

(Received February 2, 1998.)

REFERENCES

[1] J. Anděl: A remark on least-squares and naïve extrapolations in non-linear AR(1) processes. J. Forecasting *15* (1996), 549–552.

[2] P. Kall and S. W. Wallace: Stochastic Programming. Wiley, Chichester 1994.

[3] H. Tong: Non-linear Time Series. Clarendon Press, Oxford 1990.

Prof. RNDr. Jiří Anděl, DrSc. and Prof. RNDr. Václav Dupač, DrSc., Department of Probability and Statistics, Faculty of Mathematics and Physics – Charles University, Sokolovská 83, 186 00 Praha 8. Czech Republic.

e-mails: andel@karlin.mff.cuni.cz, dupac@karlin.mff.cuni.cz