

The research in this report was partially supported by the U. S. Air Force Office of Scientific Research under Contract No. AFOSR-68-1416.

THE INVERSE AUTOCORRELATIONS OF A TIME SERIES

by

William S. Cleveland

*Department of Statistics
University of North Carolina at Chapel Hill*

Institute of Statistics Mimeo Series No. 689

June, 1970

THE INVERSE AUTOCORRELATIONS OF A TIME SERIES

by

William S. Cleveland
Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27514

1. INTRODUCTION.

In this paper, the inverse autocorrelations of a time series are introduced and their use described. Let X_t for integer t be a covariance stationary time series with $EX_t = 0$ and spectral density s such that

(1) s is continuous and nonzero on $[0,1]$.

Let v_k , for $k = 0, 1, \dots$, be the autocovariances and r_k the autocorrelations of X_t so that

$$v_k = \int_0^1 e^{2\pi i k f} s(f) df$$

and $r_k = v_k/v_0$. The inverse autocovariances of X_t are defined by

$$v_k^- = \int_0^1 e^{2\pi i k f} s^{-1}(f) df$$

and the inverse autocorrelations by

$$r_k^- = v_k^-/v_0^-.$$

The research in this report was partially supported by the U. S. Air Force Office of Scientific Research under Contract No. AFOSR-68-1416.

Thus \bar{r}_k are the autocorrelations of a time series with spectral density s^{-1} .

Interest in \bar{r}_k arises from the following three properties, which will be illustrated in the remainder of the paper.

- (2) Some characteristics of X_t are better perceived in \bar{r}_k than in r_k , particularly if fitting a moving-average, autoregressive model to X_t is the goal.
- (3) The knowledge and intuition that one has of the relationship between time series and their autocorrelations will serve to interpret the inverse autocorrelations. That is, no new body of knowledge regarding the relationship between a time series and its inverse autocorrelations need be developed.
- (4) \bar{r}_k is easily estimated.

2. USING ESTIMATES OF THE INVERSE AUTOCORRELATIONS TO AID IN THE IDENTIFICATION OF A PARSIMONIOUS, MOVING-AVERAGE, AUTOREGRESSIVE MODEL.

The most common method of parametrically estimating, from a sample X_1, \dots, X_T , the mechanism generating X_t is to assume it satisfies a moving-average autoregression (which will be abbreviated to MA),

$$(5) \quad \alpha(B)X_t = \mu(B)R_t$$

where B is the backward shift operator defined by $BX_t = X_{t-1}$; R_t , the residual process, is a sequence of independent (normal) random variables with $ER_t = 0$ and $ER_t^2 = \sigma^2$; and $\alpha(B)$ and $\mu(B)$ are products of polynomials in B each of whose roots lie outside the unit circle. $\alpha(B)$ is

the autoregressive part of the model and $\mu(B)$, the moving-average part.

For example, the model might be

$$(1 + \alpha_1 B + \alpha_2 B^3) (1 + \alpha_3 B^{12}) X_t = (1 + \mu_1 B^6) R_t.$$

Having chosen the particular type of MA to be fit, the unknown parameters, that is σ^2 and the coefficients of the specified polynomials, are estimated.

Accounts of fitting these models may be found in the papers by Box and Jenkins and the paper by Hannan cited in the bibliography.

The initial step of identifying the particular type of MA to be fit is an important one. It is important to keep the number of parameters as small as possible while leaving enough to be able to get a good approximation of the truth, for each added parameter makes understanding the likelihood or least squares function more difficult.

Calculating and plotting estimates

$$\hat{r}_k = \frac{\hat{v}_k}{\hat{v}_0} = \frac{\frac{1}{T} \sum_{t=1}^{T-k} X_t X_{t+k}}{\frac{1}{T} \sum_{t=1}^T X_t^2}$$

of r_k is a good way to begin the identification procedure, for much can be said about the relationship between r_k and the coefficients of the polynomials $\alpha(B)$ and $\mu(B)$. Stralkowski and Wu (1968) have developed charts to aid in understanding this relationship, particularly for low order models.

Calculating and plotting estimates \hat{p}_k of p_k , the partial autocorrelation between X_t and X_{t-k} given $X_{t-1}, \dots, X_{t-k+1}$, has also been advocated (Box and Jenkins, 1967). One use that can be made of \hat{p}_k is to choose the value of a that would be needed if the autoregression

$$(6) \quad X_t + \sum_{j=1}^a \alpha_j X_{t-j} = R_t$$

is fit. (Any series whose spectral density satisfies (1) may be approximated by such an autoregression with a sufficiently large a .) Such a model should fit well if \hat{p}_k is nearly 0 for $k > a$. This use of \hat{p}_k will be employed in Section 3 to form estimates of r_k^- . However, other than indicating this value of a , \hat{p}_k seems to add little information beyond what is given by \hat{r}_k . Furthermore, \hat{p}_k suffers from the disadvantage that (3) does not hold for p_k . This point will be illustrated in the examples of Section 4.

That (3) holds for r_k^- can be seen by denoting the autocorrelations and inverse autocorrelations of the MA in (5) by $r_k(\alpha, \mu)$ and $r_k^-(\alpha, \mu)$ respectively. Now the spectral density of this MA is

$$s(f) = \frac{|\mu(e^{2\pi if})|^2}{|\alpha(e^{2\pi if})|^2}.$$

Thus

$$\begin{aligned} r_k^-(\alpha, \mu) &= \frac{\int_0^1 e^{2\pi i k f} s^{-1}(f) df}{\int_0^1 s^{-1}(f) df} \\ &= \frac{\int_0^1 e^{2\pi i k f} \frac{|\alpha(e^{2\pi if})|^2}{|\mu(e^{2\pi if})|^2} df}{\int_0^1 \frac{|\alpha(e^{2\pi if})|^2}{|\mu(e^{2\pi if})|^2} df} \\ &= r_k(\mu, \alpha). \end{aligned}$$

Thus in practice, estimates \hat{r}_k^- of r_k^- from the sample X_1, \dots, X_T behaving like the *autocorrelations* of an MA with autoregressive polynomial $\mu(B)$ and moving-average polynomial $\alpha(B)$, suggest the MA in (5) be fit to

X_t . Furthermore, the autocorrelation charts of Stralkowski and Wu may be used to aid in interpreting \hat{r}_k^- simply by interchanging the words "autoregressive" and "moving-average". For example, if $\hat{r}_k^- \approx r^k$ where $|r| < 1$, a first order moving-average

$$X_t = R_t + \mu_1 R_{t-1}$$

is suggested since r^k are the autocorrelations of a first order autoregression. If $\hat{r}_1^- = r$ and $\hat{r}_k^- \approx 0$ for $k > 1$ then a first order autoregression

$$X_t + \alpha_1 X_{t-1} = R_t$$

is suggested since the above values are the autocorrelations of a first order moving-average. Of course, in these simple examples \hat{r}_k^- would serve equally well to indicate the model.

Calculating and plotting estimates \hat{r}_k^- supplements the use of \hat{r}_k^- since \hat{r}_k^- can reveal effects not apparent in \hat{r}_k^- or can help to solidify opinions formed from \hat{r}_k^- about the choice of the MA model. This will be illustrated in the examples of Section 4.

3. ESTIMATION OF THE INVERSE AUTOCORRELATIONS.

If X_t is the a -th order autoregression in (6), then

$$\hat{r}_k^- = \frac{\sum_{j=0}^{a-k} \alpha_j \alpha_{j+k}}{\sum_{j=0}^a \alpha_j^2}$$

for $k = 1, \dots, a$, where $\alpha_0 = 1$, and $r_k^- = 0$ for $k > a$. As stated in Section 2, a may be chosen in practice by examining \hat{p}_k . The choice of a can also be guided by examining the variance of the fitted residuals from the model in (6); a would be chosen to be the point where the variances cease to be significantly reduced. Then if α_j are estimated from the sample X_1, \dots, X_T by $\hat{\alpha}_j$, r_k^- can be estimated by

$$\hat{r}_k^- = \frac{\sum_{j=0}^{a-k} \hat{\alpha}_j \hat{\alpha}_{j+k}}{\sum_{j=0}^a \hat{\alpha}_j^2}$$

where $\hat{\alpha}_0 = 1$. If a is chosen too small, the resulting estimates \hat{r}_k^- will be biased. If a is chosen too large (with respect to T , the number of observations) the estimates will have large variances. Thus in practice, it will sometimes be worthwhile to form \hat{r}_k^- for a few different values of a . If the \hat{r}_k^- resulting from a particular a are reliable, the values of \hat{r}_k^- should not change drastically when a is increased by a moderate amount.

Estimates $\hat{\alpha}_j$ and \hat{p}_k may be got by using the Levinson (1949, p.147) recursive formulas. This is more easily explained if the estimates of α_j , assuming X_t is an n -th order autoregression, are denoted by $\hat{\alpha}_j(n)$ for $j = 0, \dots, n$, where $\hat{\alpha}_0(n) = 1$. Then $\hat{\alpha}_j(n)$ may be calculated recursively in n using the formulas

$$\hat{\alpha}_n(n) = - \frac{\sum_{k=0}^{n-1} \hat{\alpha}_k(n-1) \hat{v}_{n-k}}{\sum_{k=0}^{n-1} \hat{\alpha}_k(n-1) \hat{v}_k}$$

and

$$\hat{\alpha}_j(n) = \hat{\alpha}_j(n-1) + \hat{\alpha}_n(n) \hat{\alpha}_{n-j}(n-1) \quad \text{for } j = 1, \dots, n-1.$$

p_k is estimated by

$$\hat{p}_k = -\hat{\alpha}_k(k).$$

$\hat{\alpha}_j$ are approximately, if the roots of $\sum_{j=0}^a \alpha_j B^j$ are not too close to the unit circle, the least squares estimates got by finding the minimum over $\alpha_1, \dots, \alpha_a$ of

$$\sum_{t=a+1}^T (X_t + \alpha_1 X_{t-1} + \dots + \alpha_a X_{t-a})^2.$$

The least squares estimates are approximately the maximum likelihood estimates if R_t is normal. Thus $\hat{r}_1^-, \dots, \hat{r}_a^-$ are approximately the least squares estimates of r_1^-, \dots, r_a^- and approximately the maximum likelihood estimates if R_t is normal, since r_1^-, \dots, r_a^- is a one-to-one function of $\alpha_1, \dots, \alpha_a$. If R_t is dangerously nonnormal, then X_t should be transformed or some modification of the least squares procedure used (Anscombe, 1967).

The remainder of this section will be devoted to the asymptotic joint distribution of $\hat{r}_1^-, \dots, \hat{r}_a^-$. It is assumed that X_t satisfies (6) where

$$(7) \quad R_t \text{ is i.i.d. with } ER_t^4 < \infty.$$

An easy computation shows that $\hat{\alpha}_j$, and therefore \hat{r}_k^- are the estimates of α_j and r_k^- described by Walker (1964, p.366, (2)). (6) and (7) insure that assumptions (1) and (5) on p.366-7 and the conditions of Theorem 2, p.371 of (Walker, 1964) hold. (Assumption (3), p.367, is not needed, as explained by Walker on p.383.) Thus Walker's Theorem 2 is applicable and implies the following results:

$$\Gamma_1^{-1/2}(\hat{r}_1^- - r_1^-, \dots, \hat{r}_a^- - r_a^-) \quad \text{and} \quad \Gamma_2^{-1/2}(\hat{\alpha}_1 - \alpha_1, \dots, \hat{\alpha}_a - \alpha_a)$$

are each asymptotically normal with mean 0 and covariance matrices Γ_1 and Γ_2 respectively; let γ_1^{jk} and γ_2^{jk} denote the k, j -th elements of Γ_1^{-1} and Γ_2^{-1} respectively and let $\alpha(B) = 1 + \alpha_1 B + \dots + \alpha_a B^a$ then

$$\gamma_2^{jk} = \int_0^1 \frac{\partial \log |\alpha(e^{2\pi i f})|^2}{\partial \alpha_k} \frac{\partial \log |\alpha(e^{2\pi i f})|^2}{\partial \alpha_j} df$$

and

$$\gamma_1^{jk} = \int_0^1 \frac{\partial \log |\alpha(e^{2\pi i f})|^2}{\partial r_k^-} \frac{\partial \log |\alpha(e^{2\pi i f})|^2}{\partial r_j^-} df.$$

An easy calculation gives

$$\gamma_2^{kj} = v_{|k-j|}.$$

This result, together with an application of the chain rule yields

$$\Gamma_2^{-1} = \Delta' \Gamma_1^{-1} \Delta$$

where the k, j -th element of Δ is

$$\frac{\partial r_k^-}{\partial \alpha_j}.$$

Thus

$$\Gamma_1 = \Delta \Gamma_2 \Delta'.$$

Γ_1 can be expressed in terms of α_j , r_k^- , and v_0^- . To simplify notation, adopt the convention $\alpha_j = 0$ for all integer j with $j < 0$ or $j > a$.

Then

$$\frac{\partial \bar{r}_k}{\partial \alpha_j} = \frac{\alpha_{j+k} + \alpha_{j-k} + 2\alpha_j \bar{r}_k}{v_0},$$

and the k, j -th element of Γ_2 is (cf. (Parzen, 1961, p.968))

$$\sum_{n=1}^a (\alpha_{k-n} \alpha_{j-n} - \alpha_{n+a-k} \alpha_{n+a-j}).$$

4. EXAMPLES.

Table 1 shows the first 18 autocorrelations, partial autocorrelations, and inverse autocorrelations of a time series. From the autocorrelations, it is clear that the model has an autoregressive part, presumably with a first order term that is causing the alternation of + and - signs. The partial autocorrelations reveal more; there is no moving-average part and 6 is the highest order autoregressive parameter needed. That is, the model

$$X_t + \sum_{j=1}^6 \alpha_j X_{t-j} = R_t$$

would adequately describe the series. The inverse autocorrelations lead also to this last conclusion but in addition suggest strongly that $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$, since \bar{r}_k are the autocorrelations of a moving-average

$$X_t = R_t + \mu_1 R_{t-1} + \mu_6 R_{t-6}.$$

The values of the table are, in fact, those of the autoregression

$$X_t + .7X_{t-1} + .4X_{t-6} = R_t.$$

Table 1

Lags	Autocorrelations								
1-9	-.84	.60	-.30	-.03	.36	-.65	.79	-.80	.68
10-18	-.46	.18	.13	-.41	.61	-.70	.67	-.54	.33
Lags	Partial Autocorrelations								
1-9	-.84	-.38	.32	-.31	.33	-.4	0	0	0
10-18	0	0	0	0	0	0	0	0	0
Lags	Inverse Autocorrelations								
1-9	.42	0	0	0	.17	.24	0	0	0
10-18	0	0	0	0	0	0	0	0	0

The inverse autocorrelation of order 5, .17, results from the 'interaction' between the first and sixth order terms of the model. Note that the symmetry between r_k and r_k^- does not exist between r_k and p_k . In this last example, p_k is not the autocorrelation function of the above moving-average.

An example of the use of the inverse autocorrelations in identifying a time series model is provided by the monthly inward station movement data in (Tiao and Thompson, 1970). The reciprocals of the original 215 observations were taken to remove the trend in the size of the seasonal oscillations. These transformed variables are the monthly average times between installations. Then a second order polynomial was fit to remove the trend in the level of the reciprocals. Let X_t be the residuals after this regression.

Table 2 gives the estimates \hat{r}_k , \hat{p}_k , and \hat{r}_k^- (with $a = 36$) for X_t . The seasonal component of order 12 is evident from the oscillations in \hat{r}_k and the high values $\hat{r}_{12} = .74$, $\hat{r}_{24} = .57$, $\hat{r}_{36} = .49$, and $\hat{r}_{48} = .43$. These last four values suggest that the seasonal component can be explained by an autoregressive term $1 + \alpha_{12} B^{12}$. The decrease in \hat{r}_{12k} is a little

Table 2: x_t

Lags	\hat{r}_k ($\hat{v}_0 = 69.9 \times 10^{-12}$)											
	1-12	.63	.35	.17	-.19	-.31	-.31	-.37	-.17	.14	.28	.6
13-24	.50	.32	.06	-.26	-.38	-.42	-.44	-.23	.03	.21	.49	.57
Lags	\hat{p}_k											
	1-12	.63	-.08	-.03	-.44	.03	-.04	-.12	.19	.30	.09	.48
13-24	-.1	-.11	-.14	-.08	-.19	-.03	-.04	-.11	.03	-.06	.02	.07
Lags	\hat{r}_k^- (a=36)											
	1-12	-.29	.06	-.12	.07	-.06	-.03	.00	.02	-.12	.20	-.16
13-24	.08	-.07	.02	-.03	.02	.11	.00	.05	-.08	.04	.05	-.05

slower than geometric, but these autocorrelations will be affected by terms of other orders. Furthermore, the inverse autocorrelations suggest $1 + \alpha_{12}B^{12}$ since $\hat{r}_{12}^- = -.21$ while \hat{r}_{24}^- and \hat{r}_{36}^- are close to 0.

The values of \hat{r}_k^- greater than .1 occur at lags 1, 3, 9-12, and 18. Some of these, however, might result from the interaction of low order components (month to month) with the very strong seasonal component. To get a cleaner look at the nonseasonal components, the model

$$(1 + \alpha_{12}B^{12})X_t = R_t$$

was fit and the fitted residuals

$$R_t^{(1)} = X_t - .74X_{t-1}$$

were analyzed.

Table 3 gives \hat{r}_k , \hat{p}_k and \hat{r}_k^- (with a=24) for $R_n^{(1)}$. In addition $\hat{r}_{36}^- = .03$ and $\hat{r}_{48}^- = .03$, which indicates the seasonal component has been

Table 3: $R_t^{(1)}$

Lags	\hat{r}_k ($\hat{v}_0 = 26.5 \times 10^{-12}$)											
1-12	.30	.23	.36	.12	.23	.21	.00	.16	.19	.05	.29	.09
12-24	.01	.17	.02	-.10	-.16	-.27	-.17	-.1	-.09	-.09	-.01	-.02
Lags	\hat{p}_k											
1-12	.30	.15	.29	-.08	.15	.03	-.12	.08	.11	-.02	.22	-.11
12-24	-.05	.01	-.04	-.19	-.26	-.17	.05	-.01	.19	.00	.15	.07
Lags	\hat{r}_k^- (9=24)											
1-12	-.09	.12	-.17	-.03	-.16	-.05	-.05	.05	-.12	.16	-.24	.07
12-24	-.04	-.11	-.08	.06	.12	.21	.04	.10	-.11	.04	-.09	-.05

adequately accounted for. \hat{r}_k^- for $k = 1, \dots, 6$ suggests that the low order components can be accounted for by an autoregressive term

$$(8) \quad \alpha_3 B^3 + \alpha_5 B^5.$$

\hat{r}_1^- has been reduced to $-.09$, whereas it was $-.29$ for X_t . There are two possibilities. One is that a first order autoregressive term $\alpha_1 B$ is not present in X_t and $\hat{r}_1^- = -.29$ arises from the interaction of other terms; the second is that the effect of a first order term in $R_t^{(1)}$ on \hat{r}_1^- is being obscured by the interaction of higher order terms. But the value $\hat{r}_1 = .30$ suggests the latter so a first order term was added to (8). The high order components in $R_t^{(1)}$, particularly those of orders 11 and 18, were ignored for the moment and the model

$$(1 + \alpha_{12} B^{12})(1 + \alpha_1 B + \alpha_3 B^3 + \alpha_5 B^5)X_t = R_t$$

was fit. The parameter estimates are $\hat{\alpha}_{12} = -.791$, $\hat{\alpha}_1 = -.188$,
 $\hat{\alpha}_3 = -.291$, and $\hat{\alpha}_5 = -.18$.

Table 4 gives the estimates \hat{r}_k , \hat{p}_k and \hat{r}_k^- (with $a=24$) for the fitted residuals $R_n^{(2)}$ from this last model. The low order components seem to have been adequately accounted for. In all three functions \hat{r}_k , \hat{p}_k and \hat{r}_k^- the two largest values occur at lags 11 and 18. Therefore, the autoregressive model with polynomial $1 + \alpha_{11}B^{11} + \alpha_{18}B^{18}$ was fit to $R_t^{(2)}$ with the result that the final model for X_t is

$$(1-.188B-.291B^3-.18B^5)(1-.791B^{12})(1-.225B^{11}+.229B^{18})X_t = R_t^{(3)}.$$

Table 4: $R_t^{(2)}$

Lags	\hat{r}_k ($v_0=19.4 \times 10^{-12}$)											
1-12	-.02	.11	.03	-.06	-.04	.00	-.06	-.06	.13	-.06	.24	-.07
13-24	.05	.16	.00	-.08	-.11	-.24	-.18	-.08	.00	.02	.10	.07
Lags	\hat{p}_k											
1-12	-.02	.11	.03	-.07	-.05	.01	-.04	-.07	.14	-.04	.21	-.08
13-24	.02	.18	.01	-.11	-.13	-.23	-.14	-.15	.08	-.01	.09	.07
Lags	\hat{r}_k^- ($a=24$)											
1-12	.09	.01	-.03	-.04	-.04	-.1	-.03	.02	-.14	.03	-.19	.11
13-24	-.02	-.14	-.02	.05	.10	.19	.13	.11	-.04	.00	-.08	-.05

Table 5 gives \hat{r}_k , \hat{p}_k and \hat{r}_k^- (with $a=24$) for the fitted residuals $R_t^{(3)}$. The estimates seem to indicate that there is at least no major disorder in the residuals.

Table 5: $R_t^{(3)}$

Lags	\hat{r}_k ($\hat{v}_0 = 11.9 \times 10^{-12}$)											
1-12	.06	.16	.03	.02	-.04	-.02	-.03	.03	.14	-.13	-.01	.03
13-24	-.03	.12	-.04	-.09	.02	-.12	-.18	-.07	.01	-.06	.18	.00
Lags	\hat{p}_k											
1-12	.06	.15	.01	-.01	-.05	-.01	-.02	.04	.15	-.16	-.04	.07
13-24	-.02	.13	-.06	-.14	.04	-.12	-.12	-.03	.04	-.03	.14	.04
Lags	\hat{r}_k^- (a=24)											
1-12	-.06	-.17	.02	-.08	.05	.05	-.02	-.07	-.11	.13	.02	-.02
13-24	.00	-.13	.06	.10	-.08	.06	.11	.02	.00	.05	-.12	-.03

5. TOPICS FOR FURTHER STUDY.

Another possible method of estimating \hat{r}_k^- would be to estimate $s(f)$ by $\hat{s}(f)$, one of the standard nonparametric estimates got by smoothing the periodogram, and then estimate \hat{r}_k^- by approximating

$$\int_0^1 \hat{s}^{-1}(f) e^{2\pi i k f} df.$$

I do not know the properties of such estimates or how they compare with the estimates of Section 3.

The definition of inverse autocorrelations can be generalized to multivariate time series. But more discussion is needed about the relationship between the correlation matrix function and multivariate, moving-average autoregressions.

BIBLIOGRAPHY.

- Anscombe, F.J. (1967) Topics in the investigation of linear relations fitted by the method of least squares. *Journal of the Royal Statistical Society (B)* **29**, 1-52.
- Box, G.E.P., Jenkins, G.M. and Bacon, D.W. (1967) Models for forecasting seasonal and non-seasonal time series. *Spectral Analysis of Time Series*, ed. B. Harris. Wiley, New York.
- Box, G.E.P. and Jenkins, G.M. (to be published) *Models for Prediction and Control*. Holden Day, New York.
- Hannan, E.J. (1969) The estimation of mixed moving average autoregressive systems. *Biometrika* **56**, 579-593.
- Levinson, N. (1949) The Wiener RMS Error Criterion in Design. Appendix B in *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* by N. Wiener. Wiley, New York.
- Parzen, E. (1961) An approach to time series analysis. *Annals of Mathematical Statistics* **32**, 951-989.
- Stralkowski, C.M. and Wu, S.M. (1968) Charts for the interpretation of low order autoregressive-moving average models. *Technical Report No. 164, Department of Statistics, University of Wisconsin.*
- Tiao, G.C. and Thompson, H.E. (1970) Analysis of telephone data. *Technical Report No. 222, Department of Statistics, University of Wisconsin.*
- Walker, A.M. (1964) Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time series. *Journal of the Australian Mathematical Society* **IV**, 363-384.