Time Series Analysis:

7. Multivariate processes

P. F. Góra

http://th-www.if.uj.edu.pl/zfs/gora/

2018

All kinds of time series taht have been discussed so far, and some of those that will be discussed in the future, have their multivariate (or vector) counterparts. For example, a process

$$\mathbf{x}_n = \mathbf{A}_1\mathbf{x}_{n-1} + \mathbf{A}_2\mathbf{x}_{n-2} + \cdots + \mathbf{A}_p\mathbf{x}_{n-p} + \mathbf{B}_0\boldsymbol{\eta}_n + \mathbf{B}_1\boldsymbol{\eta}_{n-1} + \cdots + \mathbf{B}_q\boldsymbol{\eta}_{n-q} \tag{1}$$

is a vector autoregressive, moving average process VARMA(p,q). In (1), $\mathbf{x}_n \in \mathbb{R}^m$ is a m-dimensional time series, $\mathbf{x}_{n-k}$ are its past values, $\boldsymbol{\eta}_n$ in a n-dimensional GWN, similar for its past values, and $\mathbf{A}_1, \ldots, \mathbf{A}_p, \mathbf{B}_0, \ldots, \mathbf{B}_q \in \mathbb{R}^{m\times m}$ are constant, real matrices. It is also possible to consider series in which the dimensionality of the “innovations” $\boldsymbol{\eta}$’s is different from that of the time series; in that case the matrices $\mathbf{B}_j$ are not square, but rectangular.

The need to discuss such processes arises when we observe more than one time series and we expect that they mutually influence each other.

Example

Two processes

$$x_n = \alpha_{11} x_{n-1} + \alpha_{12} y_{n-1} + \sigma_x \eta_{x,n} \quad \text{(2a)}$$

$$y_n = \alpha_{21} x_{n-1} + \alpha_{22} y_{n-1} + \sigma_y \eta_{y,n} \quad \text{(2b)}$$

together form a VAR(1) process with uncorrelated (independent) noises.

VAR(1)

For simplicity, we shall only deal with processes VAR(1), or of the type (2), or more generally,

$$\mathbf{x}_n = \mathbf{A}\mathbf{x}_{n-1} + \boldsymbol{\Sigma}\boldsymbol{\eta}_n \tag{3}$$

where $\boldsymbol{\Sigma} = \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_m\}$ meaning that the individual components of the vector noise are uncorrelated.

If the matrix $\mathbf{A}$ in (3) can be diagonalized, i.e. if there exists an invertible matrix $\mathbf{S}$ such that

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{A}_{\text{diag}} = \text{diag}\{\lambda_1, \dots, \lambda_m\} \tag{4}$$

the vector process (3) can be "diagonalized", or represented as a collection of series that no longer influence each other. Indeed, multiplying (1) by $\mathbf{S}^{-1}$ from the left, we get

$$\mathbf{z}_n = \mathbf{S}^{-1}\mathbf{x}_n = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{S}^{-1}\mathbf{x}_{n-1} + \mathbf{S}^{-1}\boldsymbol{\Sigma}\boldsymbol{\eta}_n = \mathbf{A}_{\text{diag}}\mathbf{z}_{n-1} + \mathbf{S}^{-1}\boldsymbol{\Sigma}\boldsymbol{\eta}_n \tag{5}$$

Notes

1. There still may be some interdependence between different components of $\mathbf{z}_n$ as the matrix $\mathbf{S}^{-1}\mathbf{\Sigma}$ is, in general, not diagonal and the noises acting on various components of $\mathbf{z}_n$ get correlated.
2. If the matrix $\mathbf{A}$ in (3) is not symmetrix, the "diagonalized" time series $\mathbf{z}_n$ may become *complex*.
3. For processes of higher orders VAR(p), a "diagonalization" in the spirit of Eq. (5) is possible only if all the matrices $\mathbf{A}_1, \dots, \mathbf{A}_p$ commute.

Embedding in a higher dimension

If we have a general VAR(p) process

$$\mathbf{x}_n = \mathbf{A}_1\mathbf{x}_{n-1} + \mathbf{A}_2\mathbf{x}_{n-2} + \cdots + \mathbf{A}_p\mathbf{x}_{n-p} + \mathbf{\Sigma}\boldsymbol{\eta}_n \qquad (6)$$

we can formally represent it as a VAR(1) process, but in a space of dimensionality $m \times p$. In block notation,

$$\begin{bmatrix} \mathbf{x}_n \\ \mathbf{x}_{n-1} \\ \vdots \\ \mathbf{x}_{n-p+2} \\ \mathbf{x}_{n-p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbb{I} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbb{I} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n-1} \\ \mathbf{x}_{n-2} \\ \vdots \\ \mathbf{x}_{n-p+1} \\ \mathbf{x}_{n-p} \end{bmatrix} + \mathbf{\Sigma} \begin{bmatrix} \boldsymbol{\eta}_n \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} . \qquad (7)$$

Stationarity of VAR(1)

From the “diagonalized” form of a VAR(1) process, we can clearly see that the process is stationary, if and only if all eigenvalues of the matrix $\mathbf{A}$ satisfy

$$\forall i = 1, \ldots, m\colon \quad |\lambda_i| < 1\,, \tag{8}$$

provided these eigenvalues exist. If *any* of the eigenvalues has a modulus that is greater than 1, the process is not stationary and explodes.

Note that the similarity transformation (4) and its inverse do not change the eigenvalues.

Cross-correlations

The most important quantity to analyse while dealing with multvariate series is the *cross-correlation*. Let x_n^j be the j-th component of the vector $\mathbf{x}_n$. Then

$$\rho_{jk}(l) = \frac{1}{\sigma_j \sigma_k} \left\langle \left(x_n^j - \left\langle x_n^j \right\rangle \right) \left(x_{n+l}^k - \left\langle x_n^k \right\rangle \right) \right\rangle \tag{9a}$$

where

$$\sigma_j = \sqrt{\left\langle \left(x_n^j - \left\langle x_n^j \right\rangle \right)^2 \right\rangle}\,. \tag{9b}$$

Note that $\rho_{jk}(l) \neq \rho_{kj}(l)$.

Because in practice we have only a single realization of the process at our disposal, we cannot do the statistical averaging. Therefore, instead of (9) we use

$$\left\langle x_n^j \right\rangle = \frac{1}{N} \sum_{n=1}^{N} x_n^j \tag{10a}$$

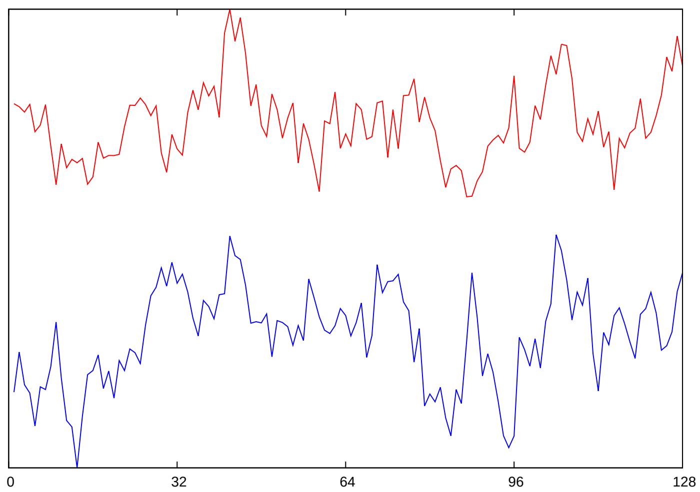
$$\sigma_j = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(x_n^j - \left\langle x_n^j \right\rangle \right)^2} \tag{10b}$$

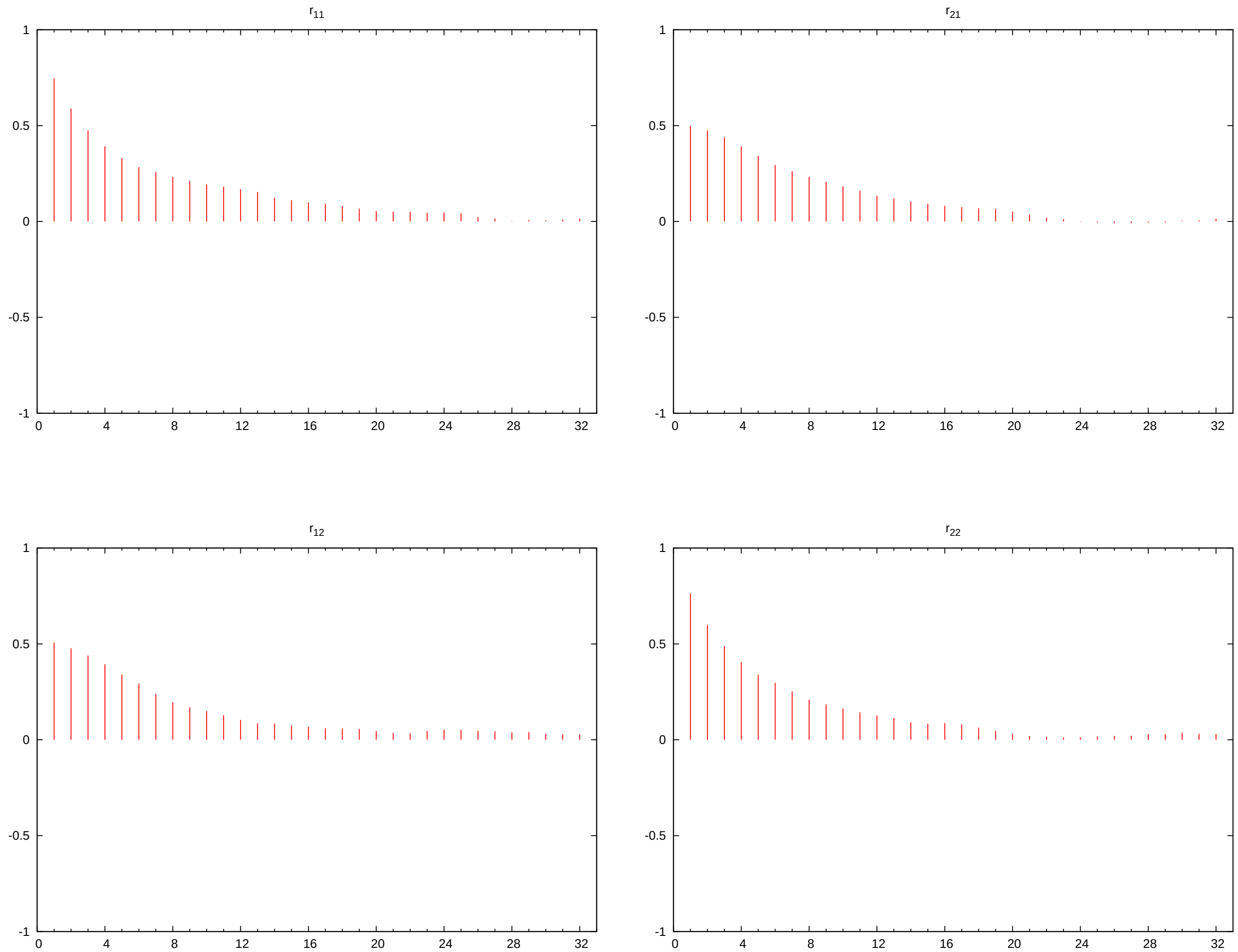
$$r_{jk}(l) = \frac{1}{(N-l)\sigma_j \sigma_k} \sum_{n=1}^{N-l} \left(x_n^j - \left\langle x_n^j \right\rangle \right) \left(x_{n+l}^k - \left\langle x_n^k \right\rangle \right) \tag{10c}$$

where N is the length of the time series.

Example 1

$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{5} \\ \frac{1}{5} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{11}$$

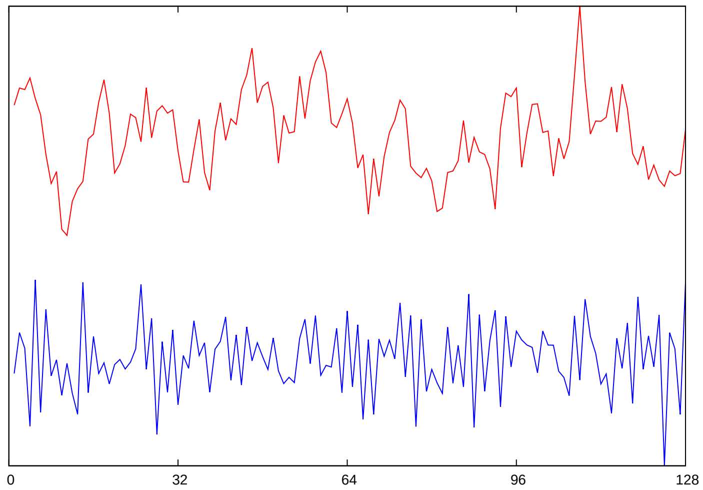


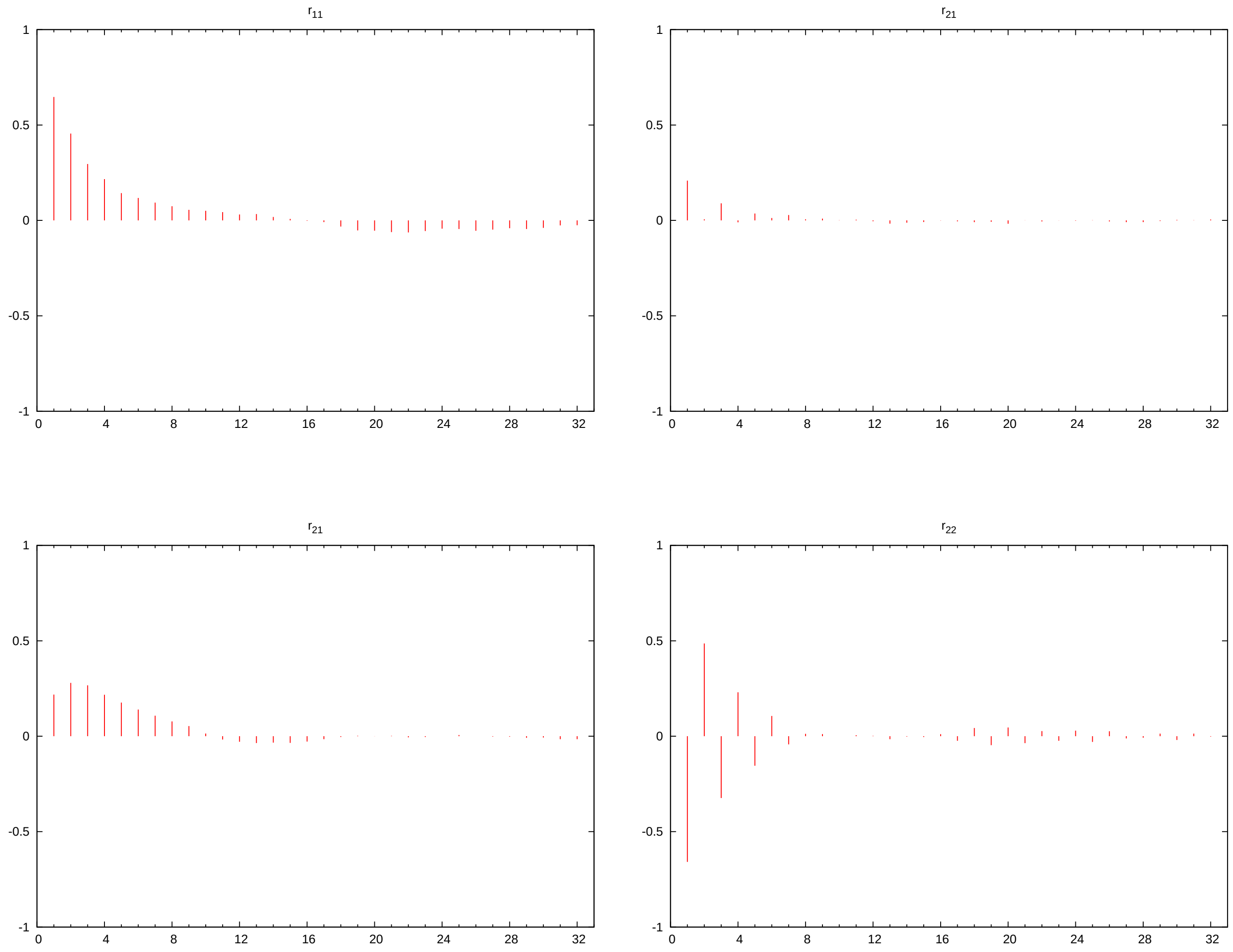


r_{11}
r_{21}
r_{12}
r_{22}
1
0.5
0
-0.5
-1
0 4 8 12 16 20 24 28 32

Example 2

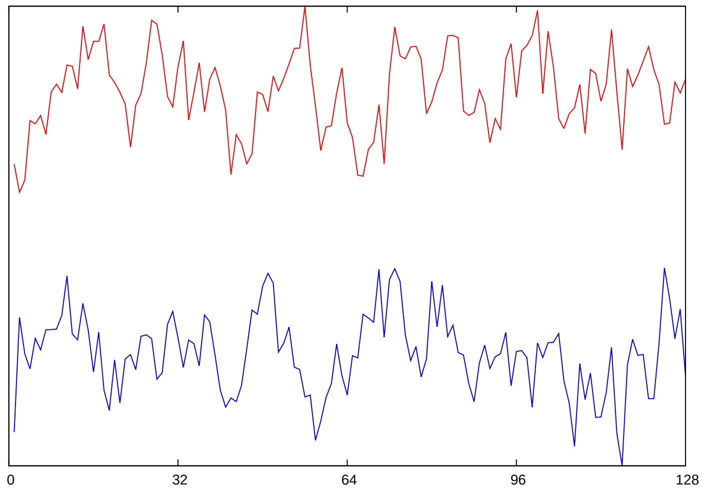
$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{5} \\ \frac{1}{5} & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{12}$$

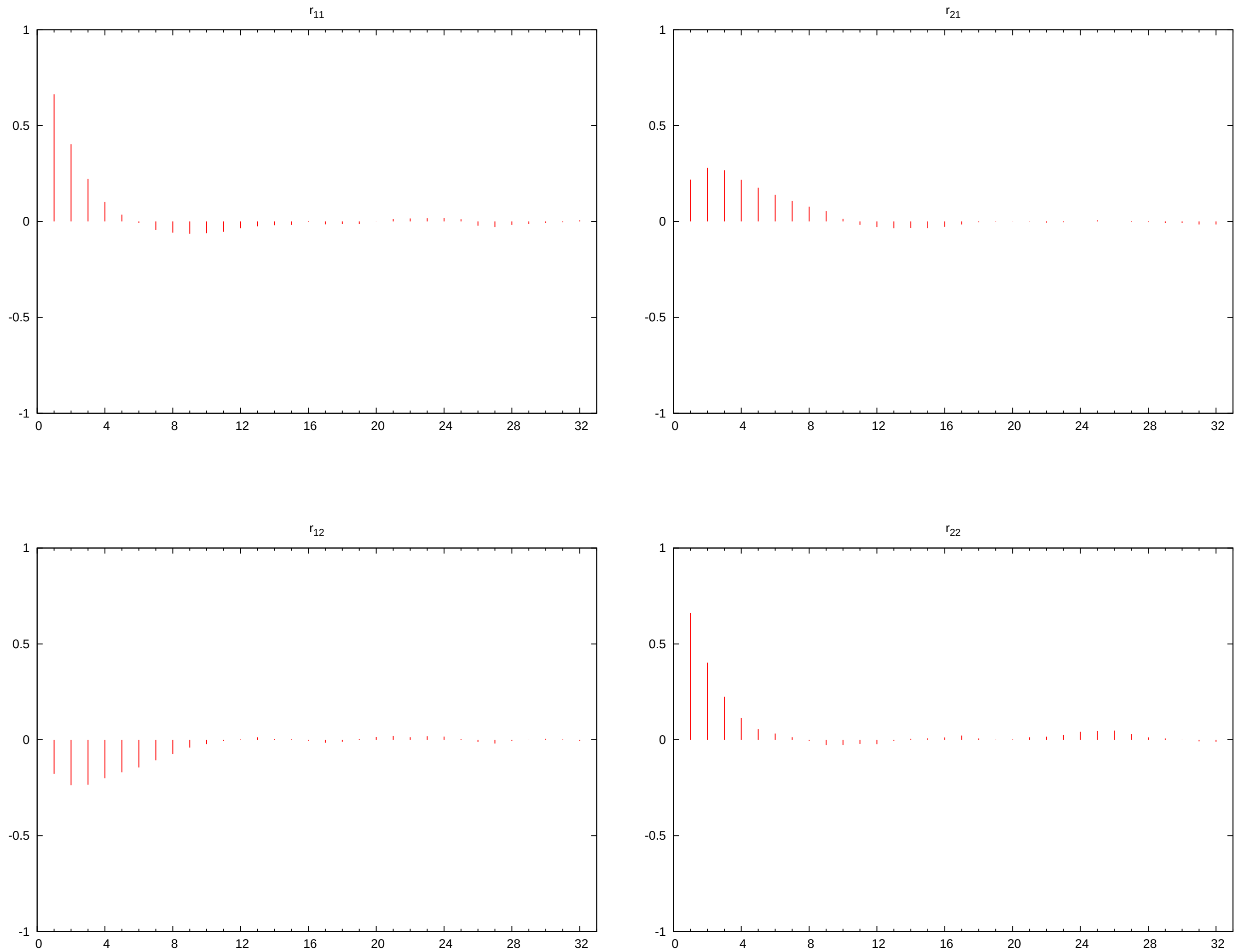



r_{11}
r_{21}
r_{21}
r_{22}

Example 3

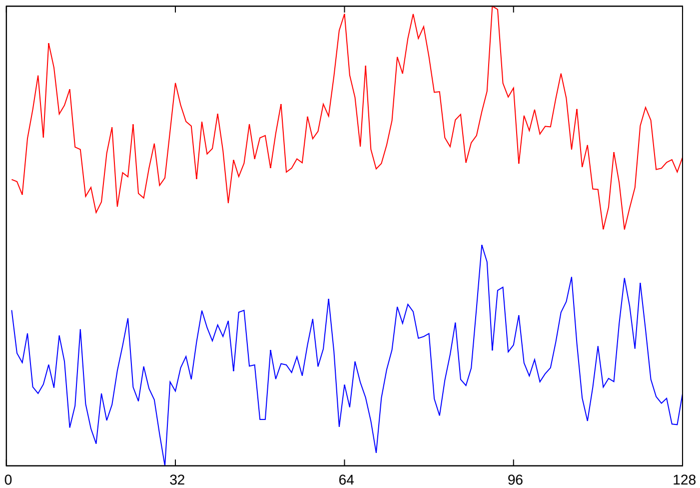
$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{5} \\ -\frac{1}{5} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{13}$$

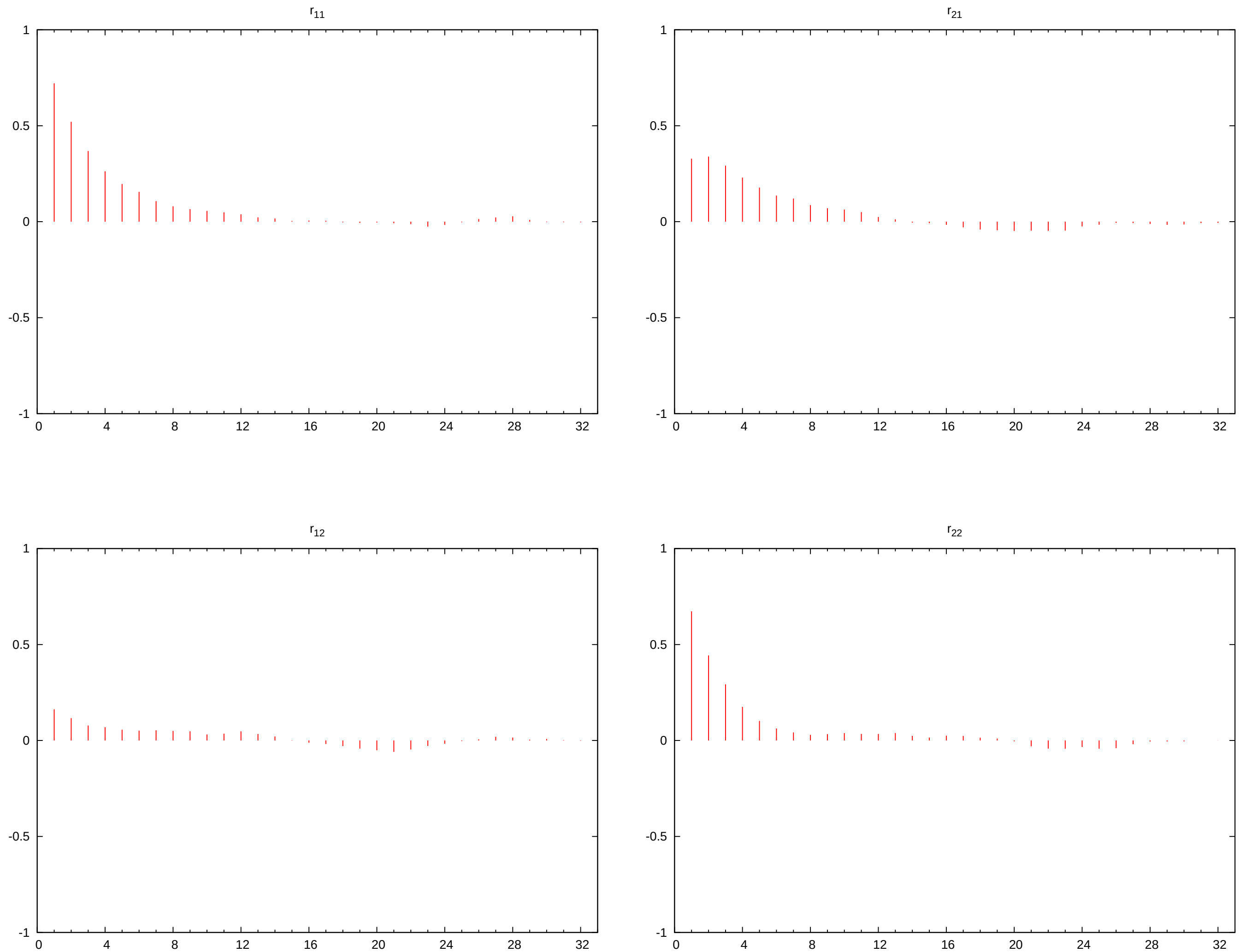



r11
r21
r12
r22
1
0.5
0
-0.5
-1
0
4
8
12
16
20
24
28
32

Example 4 — one process drives another

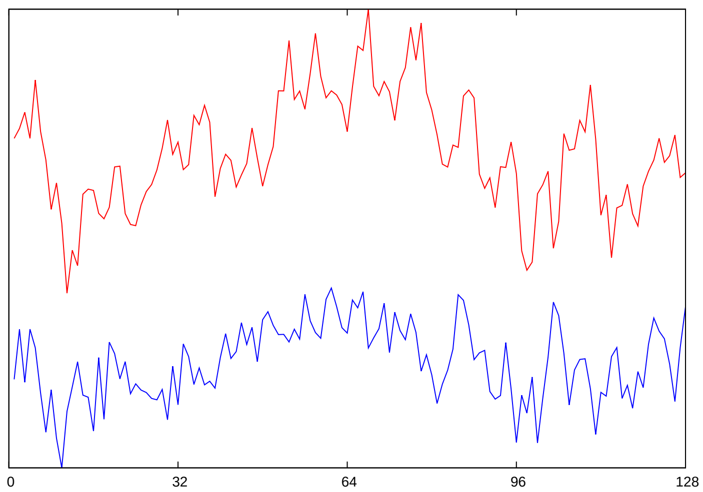
$$
\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{5} \\ 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{14}
$$

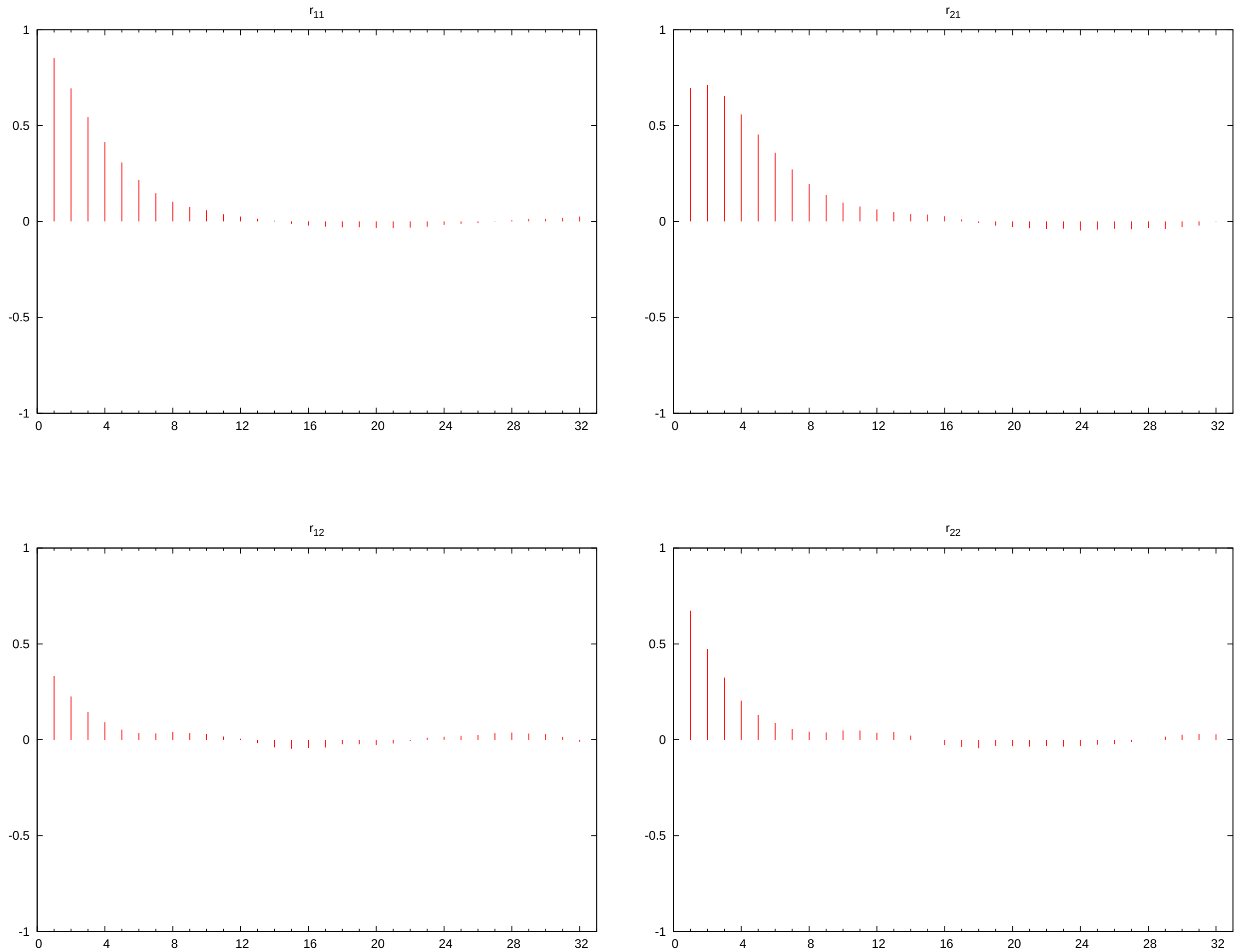



r11
r21
r12
r22
1
0.5
0
-0.5
-1
0
4
8
12
16
20
24
28
32

Example 5 — “non-diagonalizable” process

$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{15}$$

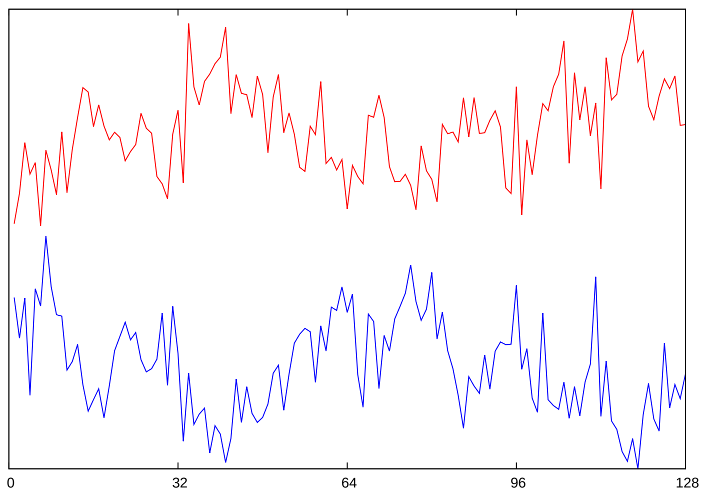


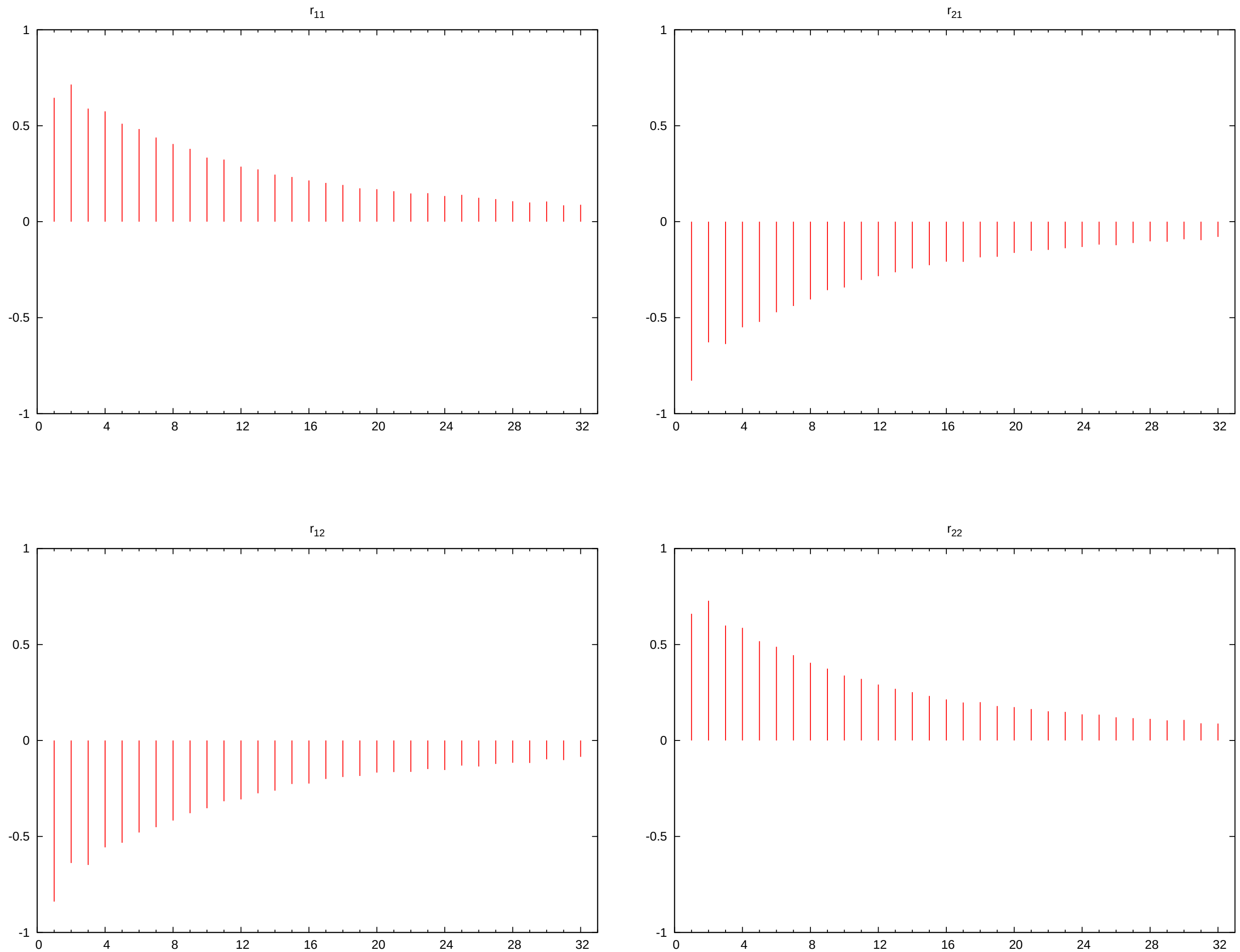


r_{11}
r_{21}
r_{12}
r_{22}

Example 6 — non-symmetric matrix, negative cross-correlations

$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{3} \\ -\frac{3}{4} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{16}$$

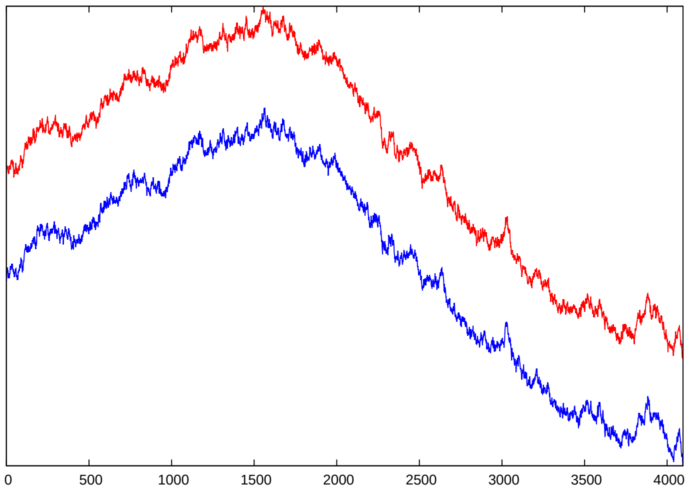




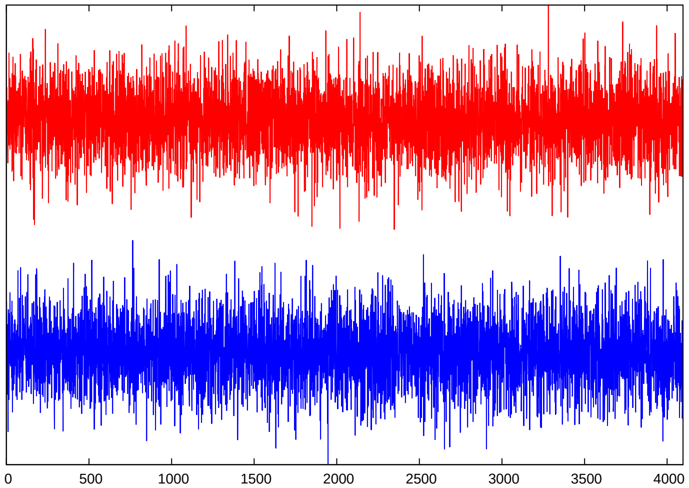
r11
r21
r12
r22
1
0.5
0
-0.5
-1
0
4
8
12
16
20
24
28
32

Example 7 — a linear trend

$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{17}$$



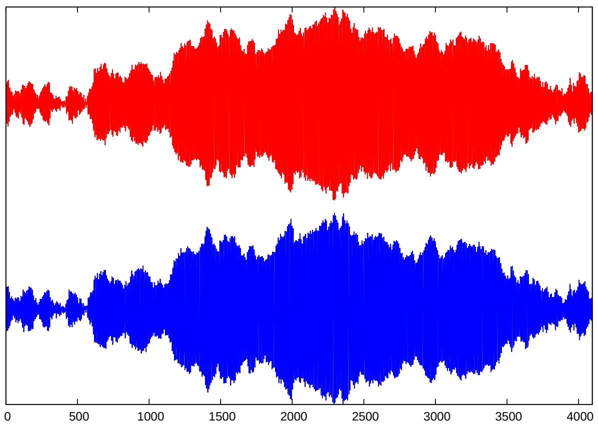
The matrix in (17) has eigenvalues $1, \frac{1}{2}$. The unit eigenvalue causes a linear trend. The series of first differences, $x^1_{n+1} - x^1_n, x^2_{n+1} - x^2_n$ are stationary.



Example 8 — another kind of nonstationarity

$$\begin{bmatrix} x_n^1 \\ x_n^2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_{n-1}^1 \\ x_{n-1}^2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \eta_n^1 \\ \eta_n^2 \end{bmatrix} \tag{18}$$

This matrix has eigenvalues $\lambda_{1,2} = \frac{1}{\sqrt{2}}(1 \pm i)$, $\left|\lambda_{1,2}\right| = 1$.



Fitting parameters to VAR(1) model

A general covariance matrix for a VAR(p) model is defined (see Eq. (9))

$$\mathbf{\Gamma}(l) = \left\langle \mathbf{x}_n \mathbf{x}_{n+l}^T \right\rangle . \tag{19}$$

For a VAR(1) we get

$$\begin{aligned} \mathbf{\Gamma}(1) = \left\langle \mathbf{x}_n \mathbf{x}_{n+1}^T \right\rangle &= \left\langle \mathbf{x}_n (\mathbf{A}\mathbf{x}_n + \mathbf{\Sigma}\boldsymbol{\eta}_n)^T \right\rangle \\ = \left\langle \mathbf{x}_n \mathbf{x}_n^T \mathbf{A}^T \right\rangle + \left\langle \mathbf{x}_n \boldsymbol{\eta}_n^T \mathbf{\Sigma}^T \right\rangle &= \left\langle \mathbf{x}_n \mathbf{x}_n^T \right\rangle \mathbf{A}^T + \left\langle \mathbf{x}_n \boldsymbol{\eta}_n^T \right\rangle \mathbf{\Sigma}^T . \end{aligned} \tag{20}$$

The last average vanishes as $\mathbf{x}_n$ does not depend on $\boldsymbol{\eta}_n$. As $\left\langle \mathbf{x}_n \mathbf{x}_n^T \right\rangle = \mathbf{\Gamma}(0)$, we get

$$\mathbf{\Gamma}(0)\mathbf{A}^T = \mathbf{\Gamma}(1) \tag{21}$$

which is a set of linear equations for the elements of $\mathbf{A}^T$. Eq. (21) is the equivalent of Yule-Walker equations for VAR(1). Estimates of $\mathbf{\Gamma}(0)$, $\mathbf{\Gamma}(1)$ can be calculated directly from the time series.

Similarly, we can calculate

$$\begin{aligned} \left\langle \mathbf{x}_{n+1}\mathbf{x}_{n+1}^T \right\rangle &= \left\langle \left(\mathbf{A}\mathbf{x}_n + \mathbf{\Sigma}\boldsymbol{\eta}_n\right)\left(\mathbf{A}\mathbf{x}_n + \mathbf{\Sigma}\boldsymbol{\eta}_n\right)^T \right\rangle \\ &= \mathbf{A}\left\langle \mathbf{x}_n\mathbf{x}_n^T \right\rangle \mathbf{A}^T + \mathbf{\Sigma}\left\langle \boldsymbol{\eta}_n\boldsymbol{\eta}_n^T \right\rangle \mathbf{\Sigma}^T \end{aligned} \tag{22}$$

As $\left\langle \boldsymbol{\eta}_n\boldsymbol{\eta}_n^T \right\rangle = \mathbb{I}$, we finally get

$$\mathbf{\Sigma}\mathbf{\Sigma}^T = \mathbf{\Gamma}(0) - \mathbf{A}\mathbf{\Gamma}(0)\mathbf{A}^T\,. \tag{23}$$

