

MULTIVARIATE TIME SERIES ANALYSIS

AN ADAPTATION OF BOX-JENKINS METHODOLOGY

Joseph N Ladalla
University of Illinois at Springfield, Springfield, IL

KEYWORDS:

Multivariate time series, Box-Jenkins ARIMA models, Principal components, Forecasting, Eigenvalues, Eigenvectors.

SUMMARY:

A time domain analysis of Multivariate Time series is suggested which exploits Box-Jenkins methodology after the given time series data is translated into the principal components based on the dispersion matrix. **After fitting suitable models for each of the principal components, these models may be converted into multivariate ARIMA models for the original data.** Similarly the forecasts for the individual principal components may be put together into a matrix of forecasts to obtain forecasts for the original multivariate time series. We apply the method on two examples - one simulated and the other real data from Gregory C. Reinsel.

As a byproduct of this approach an elegant method of determining the order of the ARIMA model is provided for multivariate time series.

1. FROM MULTIVARIATE TO UNIVARIATE ANALYSIS:

Let $\{X_t\}$ denote a multivariate stochastic process where X_t is a k-variate normal vector. Let P denote the matrix of eigen-vectors corresponding to the variance covariance matrix of X_t . Next let $\{W_t\}$ denote the principal components (vectors) of $\{X_t\}$ given by,

$$(1.1) \quad X_t P = W_t$$

Let , for the sake of simplicity,

$$(1.2) \quad W_{j,t} = \phi_{j,1} W_{j,t-1} + a_{j,t} \quad j = 1, 2, \dots, k$$

be the underlying model for W_j , $j=1, 2, \dots, k$. Here $a_{j,t}$ is assumed to have $N(0, \sigma_j^2)$ distribution.

The components of W_t viz. $W_{1,t}, \dots, W_{k,t}$ are independent univariate stochastic processes

By inverting (1.2) using (1.1) it follows that the underlying model for X_t is of the form,

$$(1.3) \quad X_t = \phi_1 X_{t-1} + a_t$$

where,

$$\phi_1 = P \lambda P^T, \text{ and } \lambda = \text{diag} \{ \phi_{1,1}, \phi_{2,1}, \dots, \phi_{k,1} \}$$

and a_t has $MVN(0, \Sigma)$ with $\Sigma = P^T D P$;
 $D = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2)$

2. MODEL ESTIMATION.

Now let $X = \{X_{1,t}, X_{2,t}, \dots, X_{k,t}\}$, $t = 1, 2, \dots, n$ denote a multivariate time series data points observed on the above stochastic process. Let P denote the matrix of sample eigen-vectors of the variance-covariance matrix of X and let $W = \{W_{1,t}, W_{2,t}, \dots, W_{k,t}\}$ denote the sample principal components given by,

$$(2.1) \quad X P = W$$

We assume that $w_{i,t}$ is a realization from $W_{i,t}$. We now use Box-Jenkins methodology to fit a model for each $w_{i,t}$. We then proceed to obtain the forecasts, say m-step ahead forecasts. Let the matrix of forecasts for W be given by,

$$(2.2) \quad W_m = \begin{matrix} & W_{11} & W_{12} & W_{13} \dots & W_{1k} \\ & W_{21} & W_{22} & W_{23} \dots & W_{2k} \\ & \dots & \dots & \dots & \dots \\ W_{m1} & W_{m2} & W_{m3} & \dots & W_{mk} \end{matrix}$$

Now from (2.1) and (2.2) we obtain the matrix of forecasts for the original time series data given by,

$$(2.3) \quad X_m = W_m P^T$$

where P^T denotes transpose of the matrix P .

Similar operations are performed to obtain say, 95% confidence intervals for the individual variates.

3. MODEL FOR THE VECTOR X_T

We illustrate hereunder how one may use the foregoing to determine a model for the vector x_t given sample data. For simplicity, let us assume that $k = 3$ and let $w_t^T = (u_t \ v_t \ w_t)$ and let, for simplicity, the fitted models for the principal components u_t , v_t and w_t be given by ARIMA (1,0,1) of the form:

$$(3.1) \quad \begin{aligned} u_t &= \gamma_1 + \lambda_1 u_{t-1} + a_t - \psi_1 a_{t-1} \\ v_t &= \gamma_2 + \lambda_2 v_{t-1} + b_t - \psi_2 b_{t-1} \\ w_t &= \gamma_3 + \lambda_3 w_{t-1} + c_t - \psi_3 c_{t-1} \end{aligned}$$

where a_t are i.i.d normal $(0, \sigma_1^2)$; b_t are i.i.d normal $(0, \sigma_2^2)$ and c_t are i.i.d normal $(0, \sigma_3^2)$

This yields a model for the vector w_t given by,

$$(3.2) \quad \begin{aligned} w_t &= \gamma + \lambda w_{t-1} + b_t + \psi b_{t-1} \\ \text{where,} \quad \gamma &= (\gamma_1 \ \gamma_2 \ \gamma_3)^T \\ b_t &= (a_t \ b_t \ c_t)^T \\ \lambda &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \psi = \begin{pmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \psi_3 \end{pmatrix} \end{aligned}$$

(3.3) Observe that b_t has normal MVN($\mathbf{0}, \mathbf{D}$) where $\mathbf{D} = \text{diag}\{\sigma_1^2, \sigma_2^2, \sigma_3^2\}$

Now (2.1) and (3.2) yield a model for x_t :

$$(3.4) \quad x_t = \delta + \phi x_{t-1} + a_t - \theta a_{t-1}$$

where,
 $\delta = P\gamma$; $\phi = P\lambda P^T$; $\theta = P\psi P^T$; $a_t = Pb_t$;

From (3.3) , (3.4) it follows that a_t is $N(\mathbf{0}, PDP^T)$

4. EXAMPLES:

Example 1.

We simulate a trivariate (normally distributed) time series data consisting of 105 observations. The data denoted by X_t , Y_t , Z_t are given in Appendix. Based on the dispersion matrix we obtain the principal components of the variables, using only the first 100 observations, leaving the last 5 observations for comparison with the forecasts. The principal components U_t , V_t and W_t are **independently** and **normally** distributed.

The matrix P of eigen-vectors is given by:

$$(4.1) \quad P = \begin{pmatrix} 0.234943 & 0.863473 & -0.446336 \\ 0.904492 & -0.362364 & -0.224914 \\ 0.355943 & 0.350865 & 0.866140 \end{pmatrix}$$

We fit appropriate ARIMA models for the principal components using Box-Jenkins methodology. Models for U_t , V_t and W_t respectively are given by,

$$(4.2) \quad \begin{aligned} u_t &= 40.22436 + 0.7985u_{t-1} + a_t; \quad a_t \sim N(0, 287.7) \\ v_t &= 35.36 + 0.37784v_{t-1} + b_t; \quad b_t \sim N(0, 144.2) \\ w_t &= -7.54 + 0.52677w_{t-1} + 0.23957w_{t-12} + c_t \\ &\quad c_t \sim N(0, 7.40) \end{aligned}$$

Equivalently,

$$(4.3) \quad \begin{aligned} w_t &= \gamma + \lambda_1 w_{t-1} + \lambda_{12} w_{t-12} + b_t \\ \text{where } w_t &= (u_t \ v_t \ w_t)^T; \\ \gamma &= (40.23436 \ 35.36 \ -7.54)^T; \\ \lambda_1 &= \text{diag} (0.7985 \ 0.37784 \ 0.52677); \\ \lambda_{12} &= \text{diag}(0 \ 0 \ 0.23957) \\ b_t &= (a_t \ b_t \ c_t)^T \text{ and } b_t \text{ is } N(\mathbf{0}, D) \\ \text{and } D &= \text{diag}(287.66 \ 144.16 \ 7.40) \end{aligned}$$

Now setting $x_t = (x_t \ y_t \ z_t)^T$, (2.1), (4.1), (4.2) and (4.3) yield the model for x_t given by,

$$(4.4) \quad \begin{aligned} x_t &= \delta + \phi_1 x_{t-1} + \phi_{12} x_{t-12} + a_t \\ \text{where, } a_t &= Pb_t; \quad a_t \text{ is } N(\mathbf{0}, \Sigma); \\ \delta &= P\gamma = (43.35 \ 25.274 \ 20.197)^T \end{aligned}$$

$$\phi_1 = P\lambda_1 P^T = \begin{pmatrix} 0.4307 & 0.1043 & -0.0224 \\ 0.1043 & 0.7295 & 0.1064 \\ -0.0224 & 0.1064 & 0.5429 \end{pmatrix}$$

$$\phi_{12} = P\lambda_{12} P^T = \begin{pmatrix} 0.0477 & 0.0240 & -0.0926 \\ 0.0240 & 0.0121 & -0.0467 \\ -0.0926 & -0.0467 & 0.1797 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 124.84 & 16.77 & 64.87 \\ & 16.77 & 254.64 \\ & & 72.84 \end{pmatrix}$$

The matrices of 5-step ahead forecasts for U, V and W, together with lower and upper 95% confidence limits are given below:

a) Forecasts:

t	U	V	W
101	173.8353	64.9940	-30.8700
102	179.0296	59.9175	-31.9553
103	183.1772	57.9993	-32.4292
104	186.4890	57.2746	-33.2898
105	189.1335	57.0007	-33.4694

b) 95% lower confidence limits :

t	U	V	W
101	140.5930	41.4616	-36.2033
102	136.4901	34.7612	-37.9833
103	135.6501	32.6197	-38.6362
104	136.0387	31.8633	-39.5456
105	136.9046	31.5849	-39.7386

c) 95% upper confidence limits:

t	U	V	W
101	207.0776	88.5265	-25.5367
102	221.5692	85.0737	-25.9273
103	230.7044	83.3789	-26.2221
104	236.9394	82.6859	-27.0340
105	241.3623	82.4166	-27.2001

Now using (2.3) and (4.1) we obtain the corresponding forecasts versus the observed values:

t	forecast	X		observed
		conf. limits lower	upper	
101	110.74	84.99	136.49	113.14
102	108.06	79.04	137.09	99.11
103	107.59	77.28	137.90	110.46
104	108.13	77.13	139.13	97.46
105	108.59	77.17	140.00	110.92

t	forecast	Y		Observed
		conf. limits lower	upper	
101	140.62	120.28	160.96	142.96
102	147.41	119.40	175.40	139.53
103	151.96	119.56	184.35	153.54
104	155.41	120.39	190.43	138.49
105	157.94	121.32	194.52	132.82

t	forecast	Z		observed
		conf. limits lower	upper	
101	57.94	33.23	82.65	63.15
102	57.07	27.88	86.26	49.63
103	57.46	26.26	88.66	53.58
104	57.64	25.35	89.93	44.87
105	58.33	25.39	91.25	52.02

The forecasts compare well with the exact..

Example 2.

Here we consider a real life example: Monthly flour price indices for three U.S. cities . (c.f Gregory C. Reinsel (1997), Table 9) We use the first 95 of the 100 observations for analysis holding the last 5 observations to compare with the forecasts.

Since the procedure is a repetition of the one adopted in the previous example, we skip details to look only at the relevant facts:

Let X = monthly flour price index for Buffalo,
 Y = monthly flour price index for Minneapolis.
 Z = monthly flour price index for Kansas city.

The matrix P of eigenvectors and the corresponding eigenvalues of the dispersion matrix of X, Y, Z are given below:

	P_1	P_2	P_3
$P =$	0.529453	0.513652	0.675160
	0.573376	0.369889	-0.731041
	0.625236	-0.774173	0.098677
eigenvalues	2425.77	43.72	6.01
Proportion of variation	0.9799	0.0177	0.0024

Using the matrices X and W for the matrices of the original time series data and the matrix of principal components, we have as in (2.1),

$$(4.5) \quad W = XP$$

Let as before, U_t , V_t , and W_t denote the first, second and third principal components. Under the assumption that the original times series are normally distributed, the principal components are independently and normally distributed. Using Box-Jenkins methodology, we fit ARIMA models to the principal components. The fitted

models for U_t , V_t and W_t are given below:

Models for U_t , V_t and W_t are given by,
(4.6)

$$U_t = 1.2363U_{t-1} - 0.2363U_{t-2} + a_t$$

$$V_t = 0.7442V_{t-1} - 0.2558V_{t-2} + b_t - 0.2064b_{t-3} + 0.2204b_{t-7}$$

$$W_t = 0.5750W_{t-1} + c_t$$

where a_t is $N(0, 184.42)$, b_t is $N(0, 8.56)$ and c_t is $N(0, 1.54)$

The model may be put in compact matrix form as,

$$(4.7) \mathbf{w}_t = \lambda_1 \mathbf{w}_{t-1} - \lambda_2 \mathbf{w}_{t-2} + \mathbf{b}_t - \gamma_3 \mathbf{b}_{t-3} + \gamma_7 \mathbf{b}_{t-7}$$

$\mathbf{b}_t$ is $N(0, \mathbf{D})$

where,

$$\lambda_1 = \text{diag} (1.2363 \quad 0.7442 \quad 0.5750)$$

$$\lambda_2 = \text{diag} (0.2363 \quad 0.2558 \quad 0)$$

$$\gamma_3 = \text{diag} (0 \quad 0.2064 \quad 0)$$

$$\gamma_7 = \text{diag} (0 \quad 0.2204 \quad 0)$$

$$\mathbf{D} = \text{diag} (184.42 \quad 8.56 \quad 1.54)$$

And as in example 1, using (4.3) and (4.5) we obtain a model for the vector $\mathbf{X}_t$ given by,

$$(4.8) \mathbf{x}_t = \phi_1 \mathbf{x}_{t-1} + \phi_2 \mathbf{x}_{t-2} + \mathbf{a}_t - \theta_3 \mathbf{a}_{t-3} - \theta_7 \mathbf{a}_{t-7};$$

and $\mathbf{a}_t$ is $N(0, \Sigma)$.

Here,

$$\phi_1 = P\lambda_1 P^T, \quad \phi_2 = P\lambda_2 P^T, \quad \theta_3 = P\gamma_3 P^T;$$

$$\theta_7 = P\gamma_7 P^T, \quad \Sigma = PDP^T$$

Accordingly, we obtain;

$$(4.9)$$

ϕ_1	=	0.8050	0.2329	0.1516
		0.2329	0.8156	0.1886
		0.1516	0.1886	0.9349
ϕ_2	=	0.1337	0.1203	-0.0235
		0.1203	0.1127	0.0115
		-0.0235	0.0115	0.2457
θ_3	=	0.0545	0.0392	-0.0821
		0.0392	0.0282	-0.0591
		-0.0821	-0.0591	0.1237
θ_7	=	0.0581	0.0419	-0.0876
		0.0419	0.0302	-0.0631
		-0.0876	-0.0631	0.1321
Σ	=	54.657	56.852	57.748
		56.852	62.624	63.551
		57.748	63.551	77.239

Observe the ease with which the model (4.8) is established using Box-Jenkins methodology

operating in the principal components space.

Finally using the forecasts for the principal components, like in example 1, we obtain the 5-step ahead forecasts together with 95% confidence intervals for the original time series given below:

t	X			observed
	95% lower limit	forecast	95% upper limit	
96	156.089	174.741	193.394	187.5
97	146.085	174.475	202.866	190.7
98	137.706	173.292	208.879	190.4
99	132.072	173.190	214.307	192.4
100	126.444	172.415	218.386	192.9

t	Y			observed
	95% lower limit	forecast	95% upper limit	
96	151.205	166.782	182.358	179.8
97	140.784	166.097	191.410	179.0
98	133.107	165.155	197.203	179.2
99	127.466	164.823	202.180	181.4
100	122.159	164.082	206.005	181.8

t	Z			observed
	95% lower limit	forecast	95% upper limit	
96	162.095	174.505	186.915	178.2
97	151.901	173.597	195.293	182.0
98	145.806	173.718	201.630	188.6
99	139.090	172.382	205.674	190.8
100	134.239	172.008	209.777	192.2

CONCLUSIONS:

By converting the original time series into principal components it appears that one can analyze the original series in the principal component space by applying Box-Jenkins methodology to the individual principal components. If this procedure is acceptable then we have an elegant method to analyze multivariate time series both to establish a model as well as to forecast.

One problem which needs mathematical justification is that the principal components are independent only up to 0-order cross-correlations. However, since identification and estimation procedures depend upon the sample acf and partial acf, which in turn depend on higher order cross correlations it is hoped that the problem does not arise.

APPENDIX:

Simulated trivariate time series data:

X	Y	Z
106.405	190.627	69.8236
103.199	186.567	64.0717
125.213	216.202	82.9996
149.28	192.67	86.1037
118.061	182.183	73.3294
135.681	178.08	79.8028
123.609	183.006	73.1775
119.67	169.606	71.5903
107.386	179.878	69.5109
103.184	155.704	59.1079
114.094	144.206	52.309
120.054	166.466	64.2461
84.621	155.456	46.125
99.35	131.834	45.9178
104.366	121.94	43.8563
125.856	109.621	49.5452
93.452	117.757	44.4521
102.997	124.693	48.3801
115.066	156.691	65.4364
128.531	132.815	66.0735
101.679	129.522	49.5604
103.118	138.485	49.2726
91.619	145.72	44.2546
106.05	142.91	51.4856
100.201	147.96	53.6439
113.581	152.793	58.5262
126.056	154.536	65.6212
108.847	165.975	58.81
97.912	136.182	47.0019
83.83	143.16	45.97
107.601	145.963	65.1901
110.36	153.876	70.4322
97.757	163.029	59.213
104.182	148.128	53.7676
114.859	157.864	58.5882
105.747	152.697	53.1915
93.911	149.361	45.1395
94.041	145.982	45.5159
98.611	138.33	48.2072
100.201	107.725	38.5165
88.707	135.912	47.7301
77.038	147.826	42.6818

X	Y	Z
88.54	154.424	50.2841
101.416	138.881	56.4294
115.772	162.512	71.8461
101.246	186.815	68.8396
114.453	192.601	79.7266
113.326	190.862	75.5322
125.198	222.856	86.5211
126.252	209.395	80.2245
120.437	201.562	70.4401
126.726	232.348	83.1272
110.813	206.013	71.6925
112.05	194.48	73.5274
108.752	190.653	65.5909
106.33	186.528	62.7813
118.326	200.918	71.8392
124.521	172.932	73.1649
124.204	172.396	74.4356
108.119	158.274	59.9918
97.963	172.891	59.4257
93.941	148.599	53.2992
89.856	153.749	50.943
85.456	144.248	46.2048
99.458	145.553	51.9204
101.673	154.751	55.9632
117.352	143.449	56.7054
117.677	140.059	52.1146
114.288	163.917	60.2398
121.491	162.065	61.0048
113.204	169.533	63.6008
100.379	167.589	54.1047
115.707	182.598	63.7947
111.279	175.07	58.8569
105.084	167.03	61.8931
107.303	177.034	66.3935
115.289	183.738	73.8751
125.418	175.354	72.9356
104.866	166.7	58.9892
99.596	139.441	44.6021
116.503	118.784	54.6768
106.761	158.826	53.6773
100.369	177.363	59.6366
115.777	196.3	71.5243
107.297	200.037	73.8049
110.316	190.499	65.213
96.653	179.838	61.2701
84.422	197.826	55.4614
114.216	205.546	70.3773
109.858	187.162	65.9143

X	Y	Z
124.751	156.648	66.1368
122.008	160.353	62.7405
125.239	197.26	75.3086
104.81	167.966	60.6377
119.74	174.405	71.4816
126.335	168.119	71.8745
140.156	139.09	76.9747
116.937	160.776	66.2836
120.667	167.939	69.1031
119.445	129.182	62.9973
113.135	142.962	63.149
99.113	139.533	49.6332
110.462	153.538	53.5748
97.458	138.493	44.8671
110.919	132.818	52.0201

REFERENCES

1. Box, G.E.P and Jenkins, G.M. (1976).
Time Series Analysis. Prentice Hall,
New Jersey.
2. Brockwell, P.J. and Davis, R.A. (1991).
Time Series: Theory and Methods.
Springer-Verlag, New York.
3. Johnson R.A., and Wichern D.W., (1992).
Applied Multivariate Statistical
Analysis. Prentice Hall, New
Jersey.