1 Simulator of an Artificial Network

In this section, a series of process simulators based on artificial networks are described (Rato and Reis 2014). For each case, a description of the process behind the simulator is given. Afterwards, the Matlab function is provided, followed by an example.

The following systems are available for simulation:

- Stationary Linear System
- Dynamic Linear System
- Stationary Non-linear System

1.1 Stationary Linear System

This system is composed by 16 nodes (or variables) causally related according to the representation provided in Figure 1. The original variables relationships were linearized according to Equation (1), where ε_i is a white noise sequence with a signal-to-noise ratio of 10 dB ($SNR = 10\log_{10}\left(\mathrm{var}(signal)/\mathrm{var}(noise)\right)$). Perturbations can be inputted to the system through a multiplicative factor (K_i) that causes a change on the model's parameters (see Equation (1)). During normal conditions, all K_i are equal to 1.

$$
\begin{aligned}
& g_8 = K_8\varepsilon_8,\ g_9 = K_9\varepsilon_9,\ g_{16} = K_{16}\varepsilon_{16} \\
& g_{10} = 1 + 0.40K_{10}g_8 + \varepsilon_{10} \\
& g_{11} = 0.56 + 0.15K_{11}g_8 + \varepsilon_{11} \\
& g_1 = 1.2K_1g_8 + 0.80g_9 + \varepsilon_1 \\
& g_2 = 0.60K_2g_1 + \varepsilon_2 \\
& g_3 = 0.05 + 0.22K_3g_1 + \varepsilon_3 \\
& g_4 = 1 + 0.4K_4g_1 + \varepsilon_4 \\
& g_5 = 0.062 + 0.16K_5g_1 + \varepsilon_5 \\
& g_6 = 0.60K_6g_1 + \varepsilon_6 \\
& g_7 = 0.70K_7g_1 + \varepsilon_7 \\
& g_{12} = 0.80K_{12}g_{16} + 0.51g_3 + \varepsilon_{12} \\
& g_{13} = 1.30K_{13}g_3 + \varepsilon_{13} \\
& g_{14} = 1 + 0.40K_{14}g_3 + \varepsilon_{14} \\
& g_{15} = 0.028 + 1.30K_{15}g_3 + \varepsilon_{15}
\end{aligned}
\tag{1}
$$

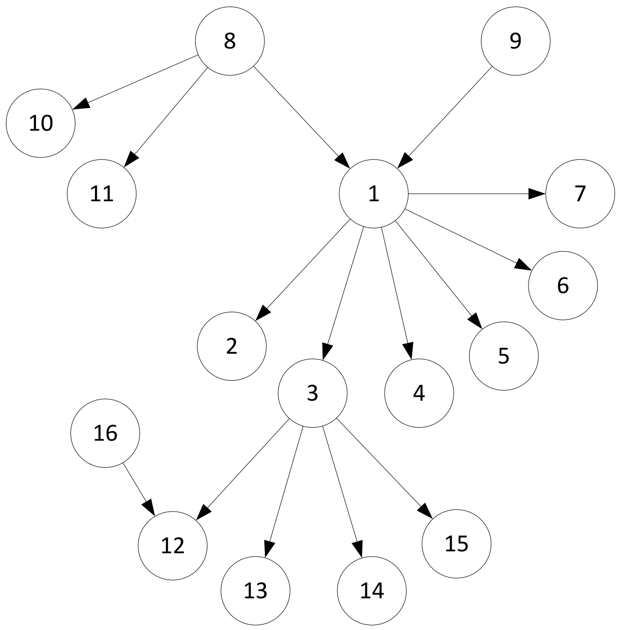

Figure 1 Graphical representation of the causal structure of the artificial network.

Matlab function `[ G ] = genenetworklin( n, K, mew )`

Inputs

n	Number of observations to simulate.
K	Multiplicative factor. Introduces deviations of K(i) in the i-th relationship, in a total of 16 relationships. The default value is ones(1,16).
mew	Variables mean value. The default value is zero (1,16).

Outputs

G	Data matrix of generated data with dimension (n × 16).

Example

```
% Generate 1000 observations with a deviation in the relationship between variables g1 and g3
n=1000;
K=ones(16,1);% Normal values;
K(3)=1.20;% Increase of 20% in the slop
mew=zeros(16,1); % Mean of the observations
[ G ] = genenetworklin( n, K, mew );
```

1.2 Dynamic Linear System

A dynamic version of the system introduced in section 1.1 can also be simulated. This system is composed by an additional multivariate time series dependency between variables according to Equation (2), where ε_i is a white noise sequence with a signal-to-noise ratio of 10 dB. Perturbation can be introduced by changing the multiplicative factor K_i.

$$
\begin{aligned}
&g_{8,t} = K_8\varepsilon_{8,t},\ g_{9,t} = K_9\varepsilon_{9,t},\ g_{16,t} = K_{16}\varepsilon_{16,t} \\
&g_{10,t} = 1+0.40K_{10}\left(g_{8,t}+0.60g_{8,t-1}-0.30g_{8,t-2}\right)+\varepsilon_{10,t} \\
&g_{11,t} = 0.56+0.15K_{11}\left(g_{8,t}+0.40g_{8,t-1}+0.60g_{8,t-2}\right)+\varepsilon_{11,t} \\
&g_{1,t} = 1.2K_1\left(g_{8,t}+0.60g_{8,t-1}+0.30g_{8,t-2}\right)+0.80g_{9,t}+\varepsilon_{1,t} \\
&g_{2,t} = 0.60K_2\left(g_{1,t}+0.50g_{1,t-1}+0.20g_{1,t-2}\right)+\varepsilon_{2,t} \\
&g_{3,t} = 0.05+0.22K_3\left(g_{1,t}-0.40g_{1,t-1}-0.20g_{1,t-2}\right)+\varepsilon_{3,t} \\
&g_{4,t} = 1+0.4K_4\left(g_{1,t}-0.20g_{1,t-1}-0.10g_{1,t-2}\right)+\varepsilon_{4,t} \\
&g_{5,t} = 0.062+0.16K_5\left(g_{1,t}+0.40g_{1,t-1}+0.60g_{1,t-2}\right)+\varepsilon_{5,t} \\
&g_{6,t} = 0.60K_6\left(g_{1,t}+0.80g_{1,t-1}+0.10g_{1,t-2}\right)+\varepsilon_{6,t} \\
&g_{7,t} = 0.70K_7\left(g_{1,t}+0.40g_{1,t-1}+0.20g_{1,t-2}\right)+\varepsilon_{7,t} \\
&g_{12,t} = 0.80K_{12}\left(g_{16,t}+0.60g_{16,t-1}+0.30g_{16,t-2}\right)+0.51g_{3,t}+\varepsilon_{12,t} \\
&g_{13,t} = 1.30K_{13}\left(g_{3,t}+0.50g_{3,t-1}+0.50g_{3,t-2}\right)+\varepsilon_{13,t} \\
&g_{14,t} = 1+0.40K_{14}\left(g_{3,t}+0.40g_{3,t-1}+0.60g_{3,t-2}\right)+\varepsilon_{14,t} \\
&g_{15,t} = 0.028+1.30K_{15}\left(g_{3,t}+0.60g_{3,t-1}-0.30g_{3,t-2}\right)+\varepsilon_{15,t}
\end{aligned}
\tag{2}
$$

Matlab function	`[ G, Gpast ] = genenetworklindyn( n, K, mew, Gpast )`		
Inputs	n	Number of observations to simulate.	
	K	Multiplicative factor. Introduces deviations of K(i) in the i-th relationship, in a total of 16 relationships. The default value is ones(1,16).	
	mew	Variables mean value. The default value is zero (1,16).	
	Gpast	Structure with information about past observations (to construct continuous data sets).	
Outputs	G	Data matrix of generated data with dimension (n × 16).	

Example

```
% Generate 1000 normal observations (Gnoc) followed by 500 observations (Gfault) with a deviation in the relationship between variables g8 and g10.
n=1000;
K=ones(16,1);% Normal values.
mew=zeros(16,1); % Mean of the observations
Gpast=[];% There are no past observations in the time series
[ Gnoc, Gpast ] = genenetworklindyn( n, K, mew, Gpast );
% Simulate contiguous fault using Gpast.
n=500;
K=ones(16,1);% Normal values;
K(10)=0.80;% Decrease of 20% in the slop
mew=zeros(16,1); % Mean of the observations
[ Gfault, Gpast ] = genenetworklindyn( n, K, mew, Gpast );
```

1.3 Stationary Non-linear System

In system the original non-linear structure of the network system of section 1.1 is approximated by polynomial relationships according to Equations (3), where ε_i is a white noise sequence with a signal-to-noise ratio of 10 dB. Perturbations are introduced by changing the multiplicative factor K_i. For normal data all K_i are equal to 1.

$$
\begin{aligned}
&g_8 = K_8\varepsilon_8,\ g_9 = K_9\varepsilon_9,\ g_{16} = K_{16}\varepsilon_{16}\\
&g_{10} = 0.020g_8^2 + 0.44K_{10}g_8 + 0.82 + \varepsilon_{10}\\
&g_{11} = -0.053K_{11}g_8^3 - 0.00068g_8^2 + 0.52g_8 + 0.50 + \varepsilon_{11}\\
&g_1 = 1.20K_1g_8 + 0.80g_9 + \varepsilon_1\\
&g_2 = 0.60K_2g_1 + \varepsilon_2\\
&g_3 = (K_3g_1 - 4)(g_1 + 4) + \varepsilon_3\\
&g_4 = 0.020K_4g_1^2 + 0.44g_1 + 0.82 + \varepsilon_4\\
&g_5 = -0.057g_1^3 - 0.077K_5g_1^2 + 0.52g_1 + 0.22 + \varepsilon_5\\
&g_6 = 0.60K_6g_1 + \varepsilon_6\\
&g_7 = 0.70K_7g_1 + \varepsilon_7\\
&g_{12} = 0.80K_{12}g_{16} + 0.60g_3 + \varepsilon_{12}\\
&g_{13} = 1.30K_{13}g_3 + \varepsilon_{13}\\
&g_{14} = 0.020K_{14}g_3^2 + 0.44g_3 - 0.82 + \varepsilon_{14}\\
&g_{15} = 1.40K_{15}g_3 + \varepsilon_{15}
\end{aligned} \tag{3}
$$

Matlab function	`[ G ] = genenetworkcub( n, K, mew )`		
Inputs	n	Number of observations to simulate.	
	K	Multiplicative factor. Introduces deviations of K(i) in the i-th relationship, in a total of 16 relationships. The default value is ones(1,16).	
	mew	Variables mean value. The default value is zero (1,16).	
Outputs	G	Data matrix of generated data with dimension (n × 16).	
Example		% Generate 1000 observations with a deviation in the relationship between variables g11 and g8 n=1000; K=ones(16,1);% Normal values; K(11)=1.20;% Increase of 20% in the third order coefficient mew=zeros(16,1); % Mean of the observations [G] = genenetworkcub(n, K, mew);	

References

Rato, T. J. and M. S. Reis (2014). "Non-causal data-driven monitoring of the process correlation structure: A comparison study with new methods." Computers & Chemical Engineering **71**: 307-322.