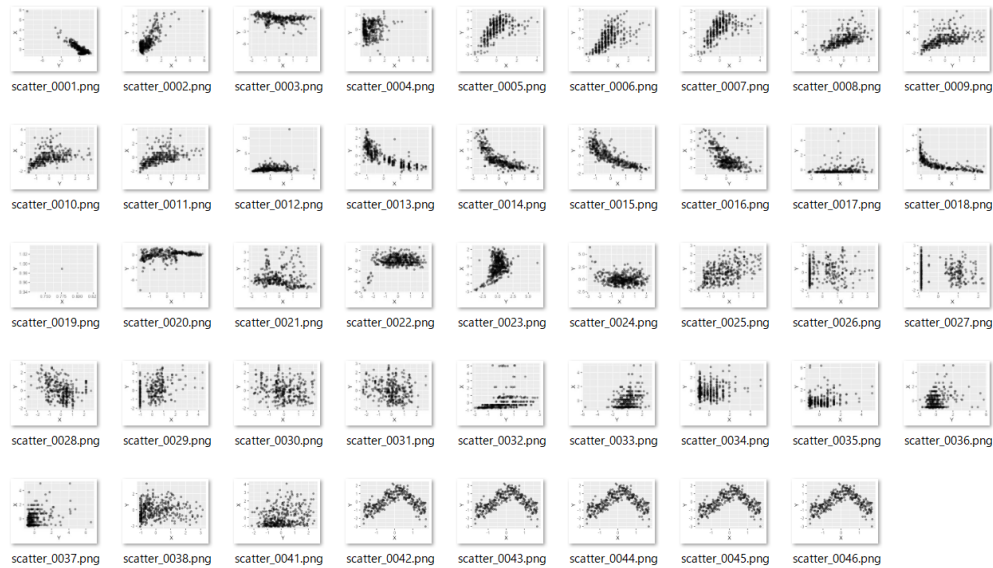


## CEP (cause-effect pairs) ベンチマーク



Database with cause-effect pairs

<https://webdav.tuebingen.mpg.de/cause-effect/README>

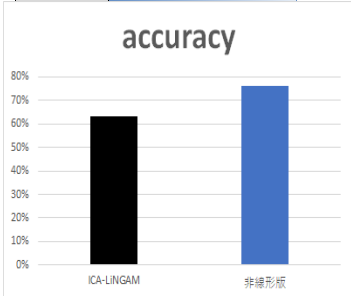
J. M. Mooij, J. Peters, D. Janzing, J. Zscheischler, B. Schoelkopf

"Distinguishing cause from effect using observational data: methods and benchmarks"

Journal of Machine Learning Research 17(32):1-102, 2016

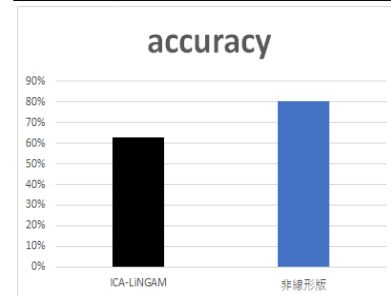
Dropout:0.5, PNL

ICA-LiNGAM	63%
非線形版	76%



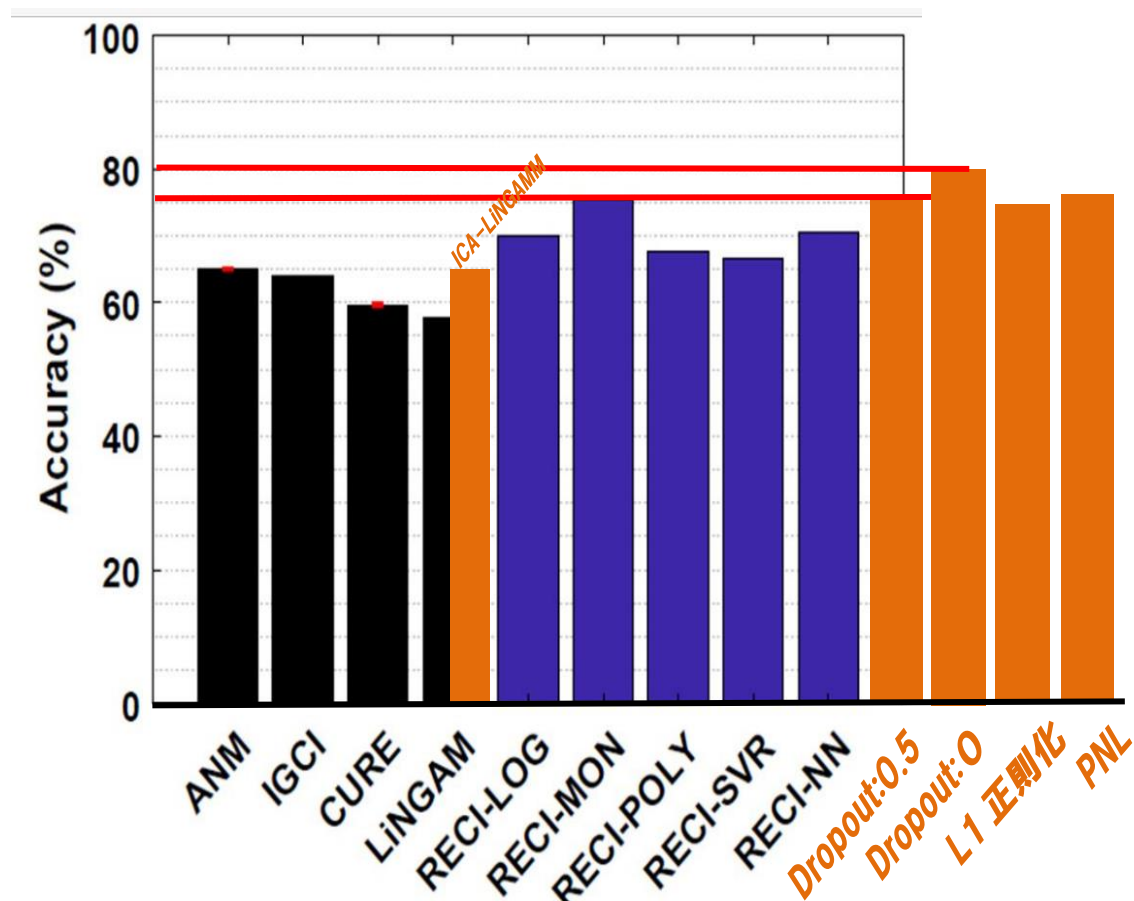
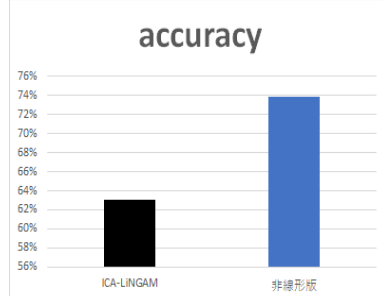
Dropout:0

ICA-LiNGAM	63%
非線形版	80%



L1 正則化

ICA-LiNGAM	63%
非線形版	74%



元グラフ引用) 非線形性に基づく大規模因果推論原理・手法の研究

<https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-17K00305/>





























































































Dropout:0.5

	X	Y	ground truth	ICA-LiNGAM	非線形
pair0001	Altitude	Temperature	->	✗	0 ✓ 1
pair0002	Altitude	Precipitation	->	✗	0 ✓ 1
pair0003	Longitude	Temperature	->	✓	1 ✓ 1
pair0004	Altitude	Sunshine hours	->	✓	1 ✓ 1
pair0005	Age	Length	->	✗	0 ✓ 1
pair0006	Age	Shell weight	->	✗	0 ✗ 0
pair0007	Age	Diameter	->	✗	0 ✓ 1
pair0008	Age	Height	->	✗	0 ✗ 0
pair0009	Age	Whole weight	->	✗	0 ✗ 0
pair0010	Age	Shucked weight	->	✗	0 ✗ 0
pair0011	Age	Viscera weight	->	✗	0 ✗ 0
pair0012	Age	Wage per hour	->	✓	1 ✓ 1
pair0013	Displacement	Fuel consumption	->	✓	1 ✓ 1
pair0014	Horse power	Fuel consumption	->	✗	0 ✗ 0
pair0015	Weight	Fuel consumption	->	✓	1 ✓ 1
pair0016	Horsepower	Acceleration	->	✓	1 ✓ 1
pair0017	Age	Dividends from stocks	->	✓	1 ✓ 1
pair0018	Age	Concentration GAG	->	✓	1 ✓ 1
pair0019	Current duration	Next interval	->	✓	1 ✓ 1
pair0020	Latitude	Temperature	->	✓	1 ✓ 1
pair0021	Longitude	Precipitation	->	✓	1 ✓ 1
pair0022	Age	Height	->	✗	0 ✓ 1
pair0023	Age	Weight	->	✗	0 ✓ 1
pair0024	Age	Heart rate	->	✗	0 ✓ 1
pair0025	Cement	Compressive strength	->	✓	1 ✓ 1
pair0026	Blast furnace	Compressive strength	->	✓	1 ✓ 1
pair0027	Fly ash	Compressive strength	->	✓	1 ✓ 1
pair0028	Water	Compressive strength	->	✗	0 ✓ 1
pair0029	Superplasticizer	Compressive strength	->	✓	1 ✓ 1
pair0030	Coarse aggregate	Compressive strength	->	✓	1 ✓ 1
pair0031	Fine aggregate	Compressive strength	->	✓	1 ✓ 1
pair0032	Age	Compressive strength	->	✗	0 ✗ 0
pair0033	Alcohol consumption	Mean corpuscular volume	->	✓	1 ✗ 0
pair0034	Alcohol consumption	Alkaline phosphatase	->	✓	1 ✓ 1
pair0035	Alcohol consumption	Alanine aminotransferase	->	✓	1 ✓ 1
pair0036	Alcohol consumption	Aspartate aminotransferase	->	✓	1 ✗ 0
pair0037	Alcohol consumption	Gamma-glutamyl transpeptidase	->	✓	1 ✗ 0
pair0038	Age	Body mass index	->	✗	0 ✓ 1
pair0039	Age	Serum insulin	->	✓	1 ✓ 1
pair0040	Age	Diastolic blood pressure	->	✓	1 ✓ 1
pair0041	Age	Plasma glucose concentration	->	✗	0 ✗ 0
pair0042	Day of the year	Temperature	->	✓	1 ✓ 1
pair0043	Temperature	Temperature at t+1	->	✓	1 ✓ 1
pair0044	Pressure	Pressure at t+1	->	✓	1 ✓ 1
pair0045	Sea level pressure	Sea level pressure at t+1	->	✓	1 ✓ 1
pair0046	Relative humidity	Relative humidity at t+1	->	✓	1 ✓ 1
				63%	76%

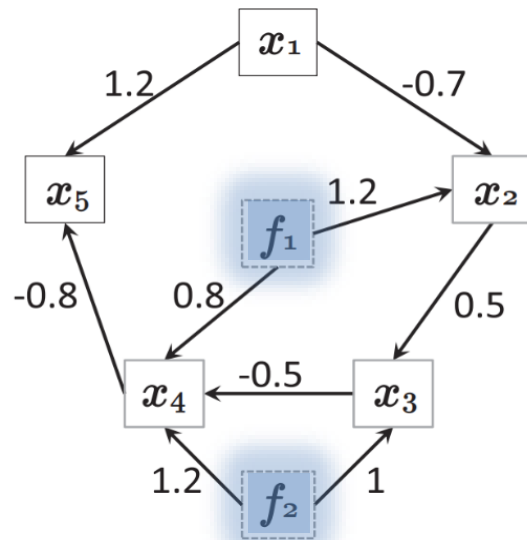
Dropout:無し

	X	Y	ground truth	ICA-LiNGAM	非線形
pair0001	Altitude	Temperature	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0002	Altitude	Precipitation	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0003	Longitude	Temperature	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0004	Altitude	Sunshine hours	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0005	Age	Length	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0006	Age	Shell weight	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0007	Age	Diameter	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0008	Age	Height	->	<div><div></div></div> 0	<div><div></div></div> 0
pair0009	Age	Whole weight	->	<div><div></div></div> 0	<div><div></div></div> 0
pair0010	Age	Shucked weight	->	<div><div></div></div> 0	<div><div></div></div> 0
pair0011	Age	Viscera weight	->	<div><div></div></div> 0	<div><div></div></div> 0
pair0012	Age	Wage per hour	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0013	Displacement	Fuel consumption	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0014	Horse power	Fuel consumption	->	<div><div></div></div> 0	<div><div></div></div> 0
pair0015	Weight	Fuel consumption	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0016	Horsepower	Acceleration	->	<div><div></div></div> 1	<div><div></div></div> 0
pair0017	Age	Dividends from stocks	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0018	Age	Concentration GAG	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0019	Current duration	Next interval	->	<div><div></div></div> 1	<div><div></div></div> 0
pair0020	Latitude	Temperature	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0021	Longitude	Precipitation	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0022	Age	Height	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0023	Age	Weight	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0024	Age	Heart rate	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0025	Cement	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0026	Blast furnace	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0027	Fly ash	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0028	Water	Compressive strength	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0029	Superplasticizer	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0030	Coarse aggregate	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0031	Fine aggregate	Compressive strength	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0032	Age	Compressive strength	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0033	Alcohol consumption	Mean corpuscular volume	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0034	Alcohol consumption	Alkaline phosphatase	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0035	Alcohol consumption	Alanine aminotransferase	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0036	Alcohol consumption	Aspartate aminotransferase	->	<div><div></div></div> 1	<div><div></div></div> 0
pair0037	Alcohol consumption	Gamma-glutamyl transpeptidase	->	<div><div></div></div> 1	<div><div></div></div> 0
pair0038	Age	Body mass index	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0039	Age	Serum insulin	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0040	Age	Diastolic blood pressure	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0041	Age	Plasma glucose concentration	->	<div><div></div></div> 0	<div><div></div></div> 1
pair0042	Day of the year	Temperature	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0043	Temperature	Temperature at t+1	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0044	Pressure at t	Pressure at t+1	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0045	Sea level pressure	Sea level pressure at t+1	->	<div><div></div></div> 1	<div><div></div></div> 1
pair0046	Relative humidity	Relative humidity at t+1	->	<div><div></div></div> 1	<div><div></div></div> 1
				63%	80%

# L1 正則化

	X	Y	ground truth	ICA-LiNGAM	非線形	
pair0001	Altitude	Temperature	->		0 	1
pair0002	Altitude	Precipitation	->		0 	1
pair0003	Longitude	Temperature	->		1 	1
pair0004	Altitude	Sunshine hours	->		1 	1
pair0005	Age	Length	->		0 	1
pair0006	Age	Shell weight	->		0 	0
pair0007	Age	Diameter	->		0 	1
pair0008	Age	Height	->		0 	0
pair0009	Age	Whole weight	->		0 	0
pair0010	Age	Shucked weight	->		0 	0
pair0011	Age	Viscera weight	->		0 	0
pair0012	Age	Wage per hour	->		1 	1
pair0013	Displacement	Fuel consumption	->		1 	1
pair0014	Horse power	Fuel consumption	->		0 	0
pair0015	Weight	Fuel consumption	->		1 	1
pair0016	Horsepower	Acceleration	->		1 	0
pair0017	Age	Dividends from stocks	->		1 	1
pair0018	Age	Concentration GAG	->		1 	0
pair0019	Current duration	Next interval	->		1 	1
pair0020	Latitude	Temperature	->		1 	1
pair0021	Longitude	Precipitation	->		1 	1
pair0022	Age	Height	->		0 	1
pair0023	Age	Weight	->		0 	1
pair0024	Age	Heart rate	->		0 	1
pair0025	Cement	Compressive strength	->		1 	1
pair0026	Blast furnace	Compressive strength	->		1 	1
pair0027	Fly ash	Compressive strength	->		1 	1
pair0028	Water	Compressive strength	->		0 	1
pair0029	Superplasticizer	Compressive strength	->		1 	1
pair0030	Coarse aggregate	Compressive strength	->		1 	1
pair0031	Fine aggregate	Compressive strength	->		1 	1
pair0032	Age	Compressive strength	->		0 	1
pair0033	Alcohol consumption	Mean corpuscular volume	->		1 	0
pair0034	Alcohol consumption	Alkaline phosphatase	->		1 	1
pair0035	Alcohol consumption	Alanine aminotransferase	->		1 	1
pair0036	Alcohol consumption	Aspartate aminotransferase	->		1 	0
pair0037	Alcohol consumption	Gamma-glutamyl transpeptidase	->		1 	0
pair0038	Age	Body mass index	->		0 	1
pair0039	Age	Serum insulin	->		1 	1
pair0040	Age	Diastolic blood pressure	->		1 	1
pair0041	Age	Plasma glucose concentration	->		0 	0
pair0042	Day of the year	Temperature	->		1 	1
pair0043	Temperature at t	Temperature at t+1	->		1 	1
pair0044	Pressure at t	Pressure at t+1	->		1 	1
pair0045	Sea level pressure at t	Sea level pressure at t+1	->		1 	1
pair0046	Relative humidity at t	Relative humidity at t+1	->		1 	1
				63%	74%	

LiNGAM\_latest3c.csv



f1, f2 が交絡変数

テストは f1, f2 もデータにある場合と f1, f2 がデータに無い場合をテストする。

※Pairwise LiNGAM. で検証に使われたデータ

$$\text{accuracy} = \frac{\text{正解エッジ数}}{\text{有効全エッジ数}}$$

$$\text{点数} = \text{正解エッジ点}(1) + \text{不正解エッジ}(-1)$$

線形因果探索は ICA-LiNGAM を適用し、因果効果が 0.01 未満は削除

非線形回帰のため係数はなく代わりに feature importance を表記

※feature importance が 0.7 以下のエッジは削除

非線形モデル	accuracy
LiNGAM_latest3.csv	0.625
nonlinear_LiNGAM_latest3c.csv	0.7143
nonlinear_LiNGAM_latest3b.csv	0.7143
nonlinear_LiNGAM_latest3a.csv	0.7143
nonlinear.csv	1
nonlinear2.csv	0.8
平均	0.76132



線形モデル ICA-LiNGAM	accuracy
LiNGAM_latest3.csv	0.625
nonlinear_LiNGAM_latest3c.csv	0.5
nonlinear_LiNGAM_latest3b.csv	0.5
nonlinear_LiNGAM_latest3a.csv	0.5556
nonlinear.csv	0.75
nonlinear2.csv	0.6667
平均	0.59955

非線形モデル	点数
LiNGAM_latest3.csv	2
nonlinear_LiNGAM_latest3c.csv	3
nonlinear_LiNGAM_latest3b.csv	3
nonlinear_LiNGAM_latest3a.csv	3
nonlinear.csv	3
nonlinear2.csv	3
平均	2.83333



線形モデル ICA-LiNGAM	点数
LiNGAM_latest3.csv	2
nonlinear_LiNGAM_latest3c.csv	0
nonlinear_LiNGAM_latest3b.csv	0
nonlinear_LiNGAM_latest3a.csv	1
nonlinear.csv	2
nonlinear2.csv	2
平均	1.16667

データは線形、アルゴリズムは線形を仮定

## 線形因果構造

LiNGAM\_latest3.csv

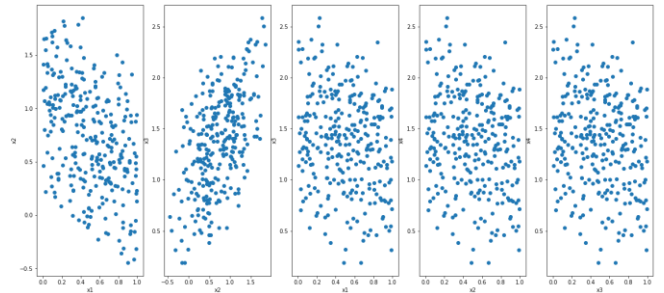
$$x_1 = e_1$$

$$x_2 = -0.7x_1 + 1.2f_1 + e_2$$

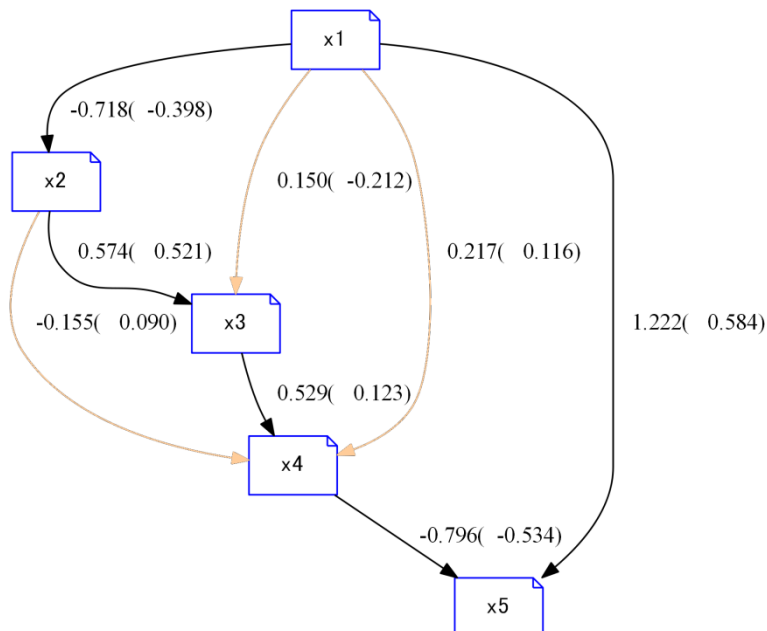
$$x_3 = 0.5x_2 + f_2 + e_3$$

$$x_4 = -0.5x_3 + 0.8f_1 + 1.2f_2 + e_4$$

$$x_5 = 1.2x_1 - 0.8x_4 + e_5$$



因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	3	-1	-3
合計	8	0	2
accuracy	62.50%		



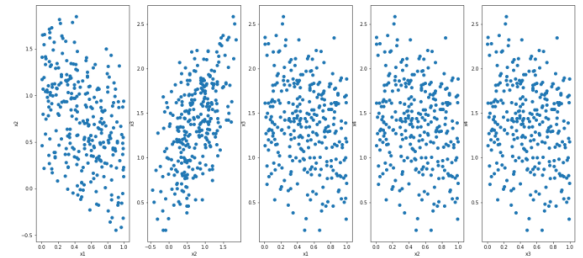


データは線形、アルゴリズムは線形を仮定しない

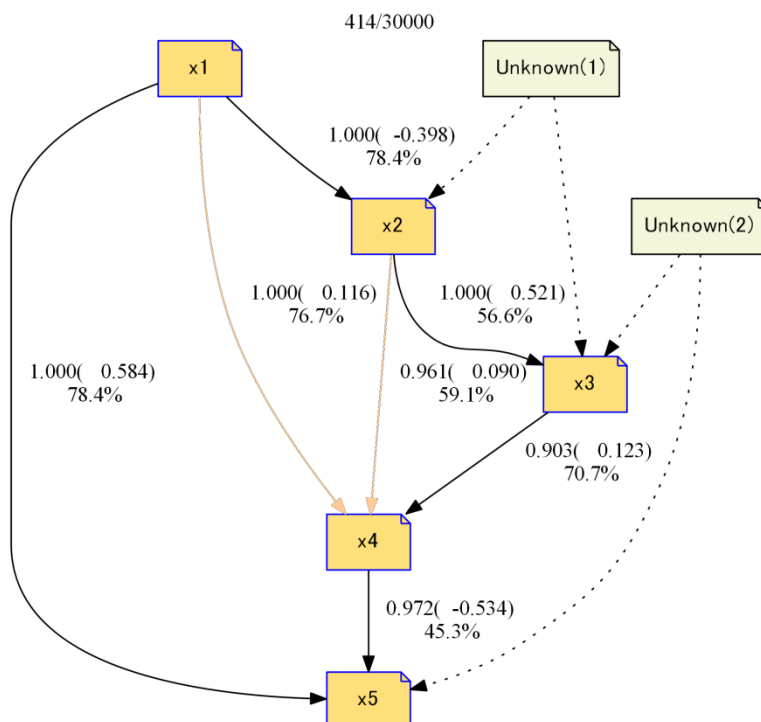
## 線形因果構造

LiNGAM\_latest3.csv

$$\begin{aligned}x_1 &= e_1 \\x_2 &= -0.7x_1 + 1.2f_1 + e_2 \\x_3 &= 0.5x_2 + f_2 + e_3 \\x_4 &= -0.5x_3 + 0.8f_1 + 1.2f_2 + e_4 \\x_5 &= 1.2x_1 - 0.8x_4 + e_5\end{aligned}$$



因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	2	-1	-2
合計	7	0	3
<b>accuracy</b>	<b>71.43%</b>		

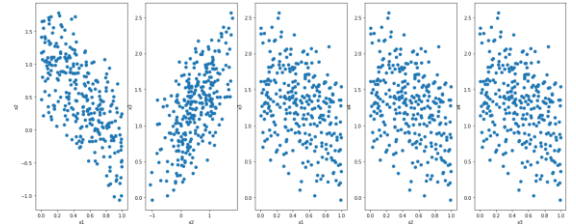


データは一部非線形、アルゴリズムは線形を仮定しない

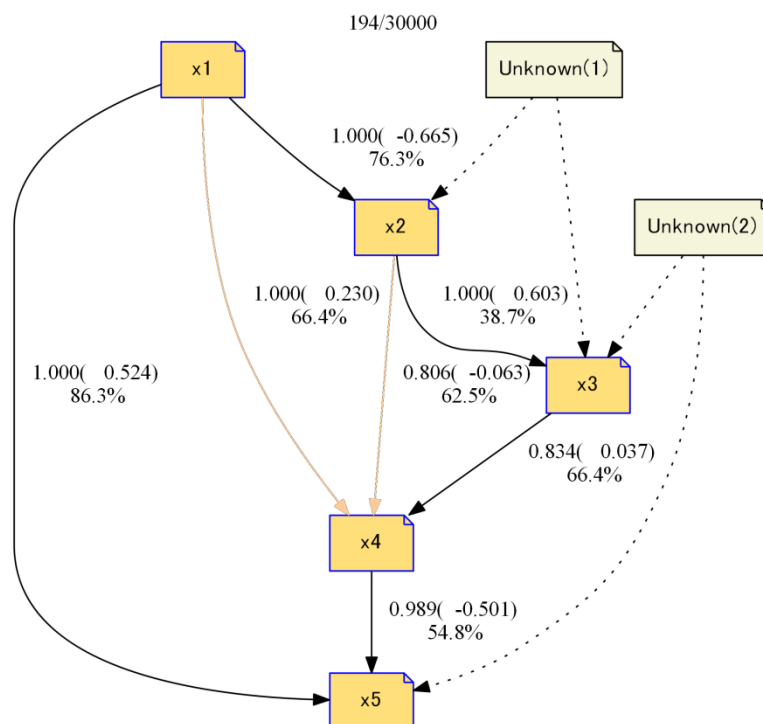
## 非線形因果構造

nonlinear\_LiNGAM\_latest3c.csv

$$\begin{aligned}x_1 &= e_1 \\x_2 &= -0.7x_1(1 + x_1) + 1.2f_1 + e_2 \\x_3 &= 0.5x_2 + f_2 + e_3 \\x_4 &= -0.5x_3 + 0.8f_1 + 1.2f_2 + e_4 \\x_5 &= 1.2x_1 - 0.8x_4 + e_5\end{aligned}$$



因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	2	-1	-2
合計	7	0	3
accuracy	71.43%		



データは一部非線形、アルゴリズムは線形を仮定しない

## 非線形因果構造

nonlinear\_LiNGAM\_latest3b.csv

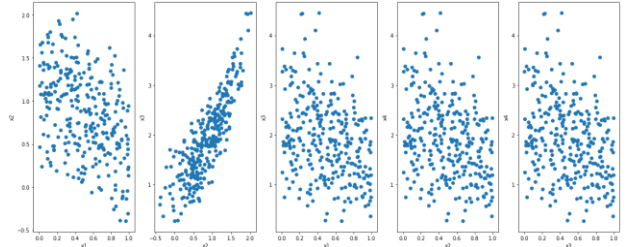
$$x_1 = e_1$$

$$x_2 = -0.7x_1x_1 + 1.2f_1 + e_2$$

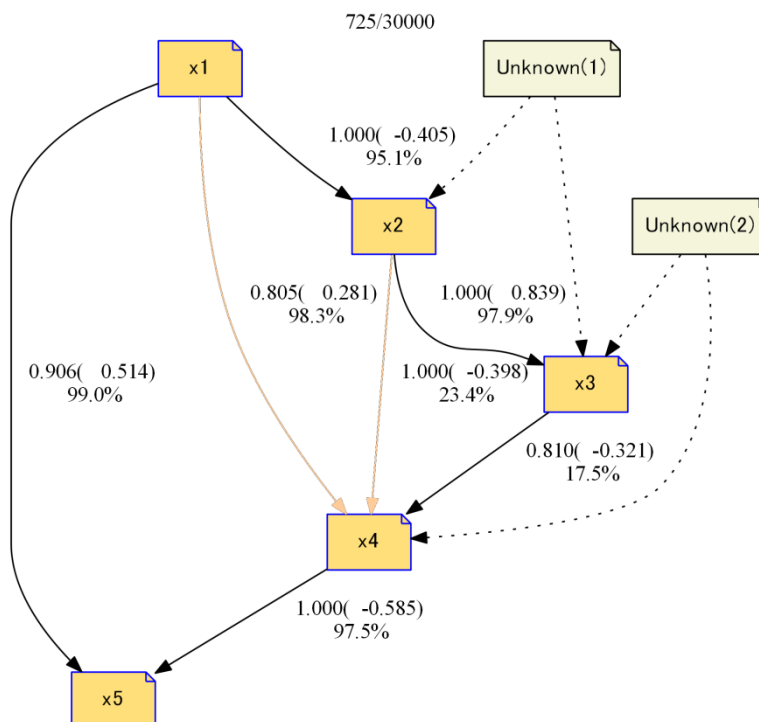
$$x_3 = 0.5x_2(1 + x_2) + f_2 + e_3$$

$$x_4 = -0.5x_3 + 0.8f_1 + 1.2f_2 + e_4$$

$$x_5 = 1.2x_1(1 + x_1) - 0.8x_4(1 + x_4) + e_5$$



因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	2	-1	-2
合計	7	0	3
accuracy	71.43%		



データは全部非線形、アルゴリズムは線形を仮定しない

## 非線形因果構造

nonlinear\_LiNGAM\_latest3a.csv

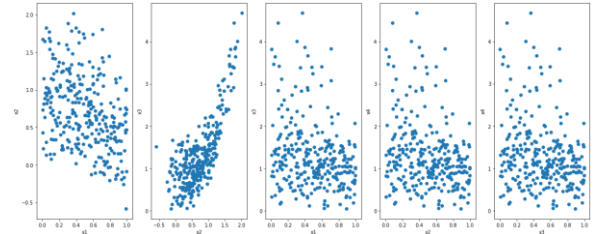
$$x_1 = e_1$$

$$x_2 = -0.7x_1^2 + 1.2f_1^2 + e_2$$

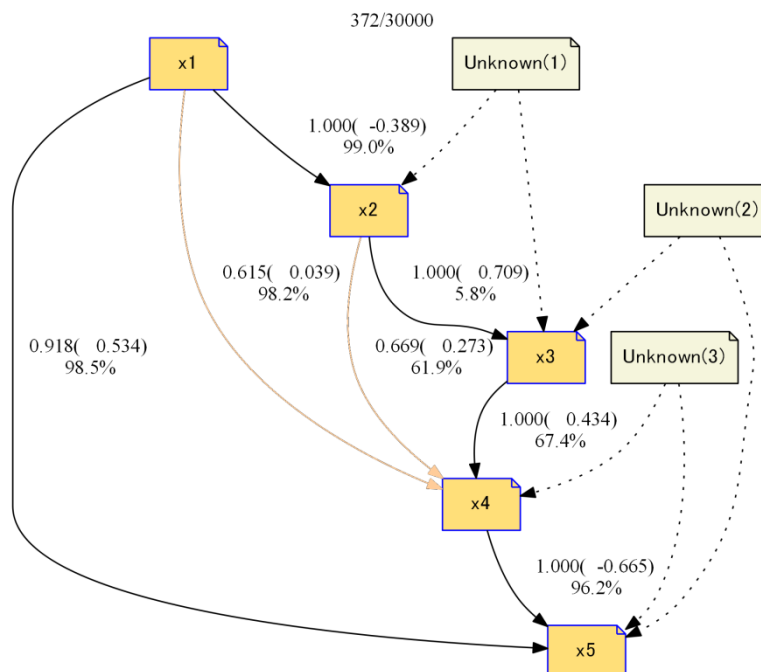
$$x_3 = 0.5x_2^3 + f_2^2 + e_3$$

$$x_4 = -0.5\sqrt{|x_3|} + (0.8f_1 + 1.2f_2)^2 + e_4$$

$$x_5 = 1.2x_1^2 - 0.8\log(|x_4| + 1)^2 + e_5$$



因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	2	-1	-2
合計	7	0	3
accuracy	71.43%		



データは全部非線形、アルゴリズムは線形を仮定しない

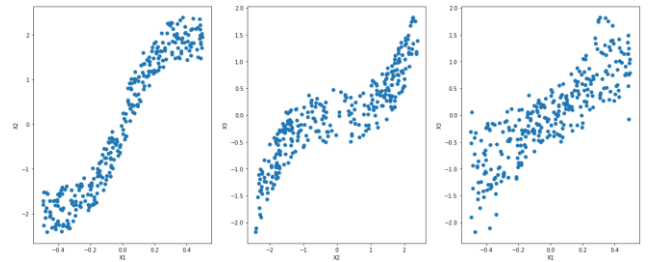
## 非線形因果構造

nonlinear.csv

$$x_1 = e_1$$

$$x_2 = 2 \tanh(5x_1) + e_2$$

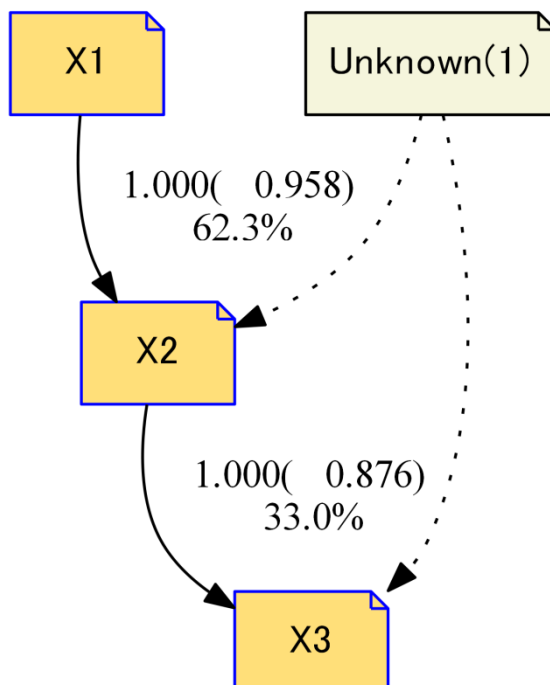
$$x_3 = \left(\frac{x_2}{2}\right)^3 + e_3$$



Causal Discovery with Cascade Nonlinear Additive Noise

因果方向	数	点数	点数
正しい因果	3	1	3
正しくない因果	0	-1	0
合計	3	0	3
accuracy	100.00%		

27/30000



データは全部非線形、アルゴリズムは線形を仮定しない

## 非線形因果構造

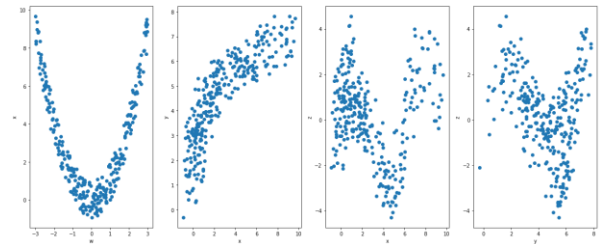
Nonlinear2.csv

$$x_1 = e_1$$

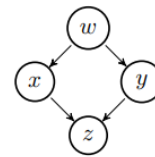
$$x_2 = x_1^2 + e_2$$

$$x_3 = 4\sqrt{|e_1|} + e_3$$

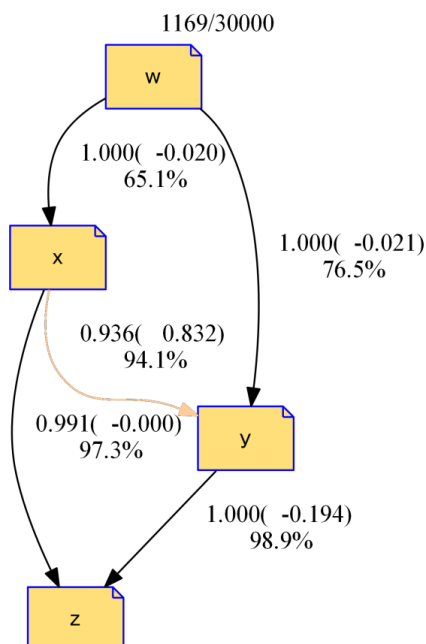
$$x_4 = 2 \sin(x_2) + 2 \sin(x_3) + e_4$$



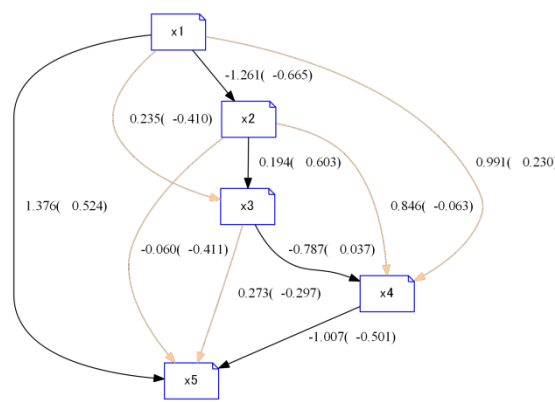
We also did experiments for 4 variables  $w, x, y$  and  $z$  with a diamond-like causal structure. We took  $w \sim U(-3, 3)$ ,  $x = w^2 + n_x$  with  $n_x \sim U(-1, 1)$ ,  $y = 4\sqrt{|w|} + n_y$  with  $n_y \sim U(-1, 1)$ ,  $z = 2 \sin x + 2 \sin y + n_z$  with  $n_z \sim U(-1, 1)$ . We sampled 500  $(w, x, y, z)$  tuples from the model and applied the algorithm described in Section 4 in order to reconstruct the DAG structure. The simplest DAG that was consistent with the data (with significance level 2% for each test) turned out to be precisely the true causal structure. All five other DAGs for which the true DAG is a subgraph were also consistent with the data.



因果方向	数	点数	点数
正しい因果	4	1	4
正しくない因果	1	-1	-1
合計	5	0	3
accuracy	80.00%		

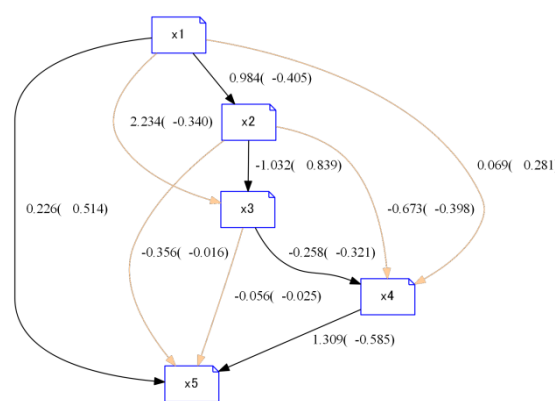


ICA-LiNGAM



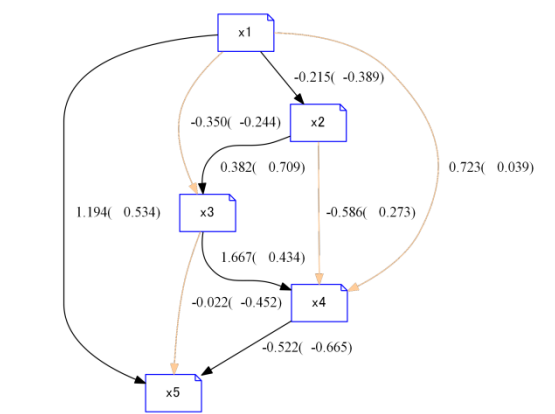
nonlinear\_LiNGAM\_latest3c.csv

因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	5	-1	-5
合計	10	0	0
accuracy	50.00%		



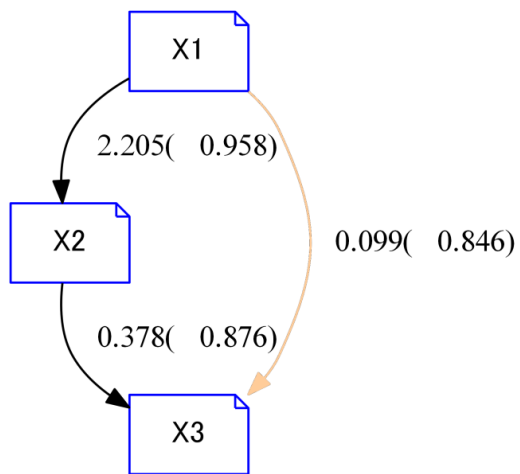
nonlinear\_LiNGAM\_latest3b.csv

因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	5	-1	-5
合計	10	0	0
accuracy	50.00%		



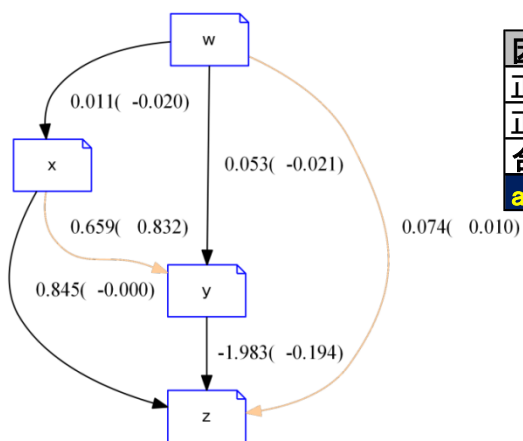
nonlinear\_LiNGAM\_latest3a.csv

因果方向	数	点数	点数
正しい因果	5	1	5
正しくない因果	4	-1	-4
合計	9	0	1
accuracy	55.56%		



nonlinear.csv

因果方向	数	点数	点数
正しい因果	3	1	3
正しくない因果	1	-1	-1
合計	4	0	2
<b>accuracy</b>	<b>75.00%</b>		



nonlinear2.csv

因果方向	数	点数	点数
正しい因果	4	1	4
正しくない因果	2	-1	-2
合計	6	0	2
<b>accuracy</b>	<b>66.67%</b>		



$$g(x_i) = f_i(\mathbf{pa}(x_i)) + \epsilon_i$$

$\mathbf{pa}(x_i)$ は $x_i$ の親（原因）となる変数、 $g_i, f_i$ は非線形関数  $f_i$  が線形、 $g_i$  が恒等関数の場合は通常の LiNGAM（線形モデル）になる。親（原因）となる変数を見つければ良いが実験モデルとして以下の B 行列を想定する。B 行列は非巡回の因果関係であれば下半三角行列になるため変数の置換を与えれば任意の非巡回の因果関係を与える事が出来る。

非巡回の因果関係であれば以下の行列を定義できる。 $B_{ij} = 1$ 。

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ B_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ B_{n1} & B_{n2} & \cdots & 0 \end{pmatrix}$$

$$B_{ij} \neq 0$$

行方向が子（結果）で列方向が親（原因）として与える事が出来る。

未観測の潜在共通変数が存在する可能性が考えられる場合、変数 $x_l$ に対して $\mu(x_l)$ を指定した以下のモ

デルを定義する。

$$g(x_i) = f_i(\mathbf{pa}(x_i), \boldsymbol{\mu}(x_i)) + \epsilon_i$$

一般化ガウス分布

$$\boldsymbol{\mu} \sim \text{Generalized\_Gaussian}(\beta, \rho, \tilde{x})$$

$$\equiv \frac{\beta^{\frac{1}{2}}}{2\Gamma(1 + \frac{1}{\rho})} \exp\left(-\beta^{\frac{1}{2}}|x - \tilde{x}|^{\rho}\right)$$

$\beta, \rho$ は未定のパラメータで以下の最適化によって定める。

$$mse_i = MSE(x_i)$$

$$= \frac{1}{n} \sum (g(x_i) - f_i(\mathbf{pa}(x_i), \mu(x_i)))^2$$

$$MI(\mathbf{e}_i, \mathbf{e}_j) \equiv \iint p(\mathbf{e}_i, \mathbf{e}_j) \log \left( \frac{p(\mathbf{e}_i, \mathbf{e}_j)}{p(\mathbf{e}_i)p(\mathbf{e}_j)} \right) d\mathbf{e}_i d\mathbf{e}_j$$

$MI$ は相互情報量で0であれば独立変数になる。つまり、 $\mu$ は残差と相互情報量がともに最小になるように未定のパラメータ $\beta, \rho$ を求める。この最適化問題は多目的最適化となるため以下の損失関数を最小化する。

拡大 Tchebyshev・スカラー化関数

$$\begin{aligned} \textit{loss} = \max & \left( w1 \max(\mathbf{e}_i), \quad w2 \max(MI(\mathbf{e}_i, \mathbf{e}_j)) \right) \\ & + \varepsilon \left( w1 \max(\mathbf{e}_i) + w2 \max(MI(\mathbf{e}_i, \mathbf{e}_j)) \right) \end{aligned}$$

$$\varepsilon = 0.0001$$

$$w1 = 0.7$$

$$w2 = 0.4$$