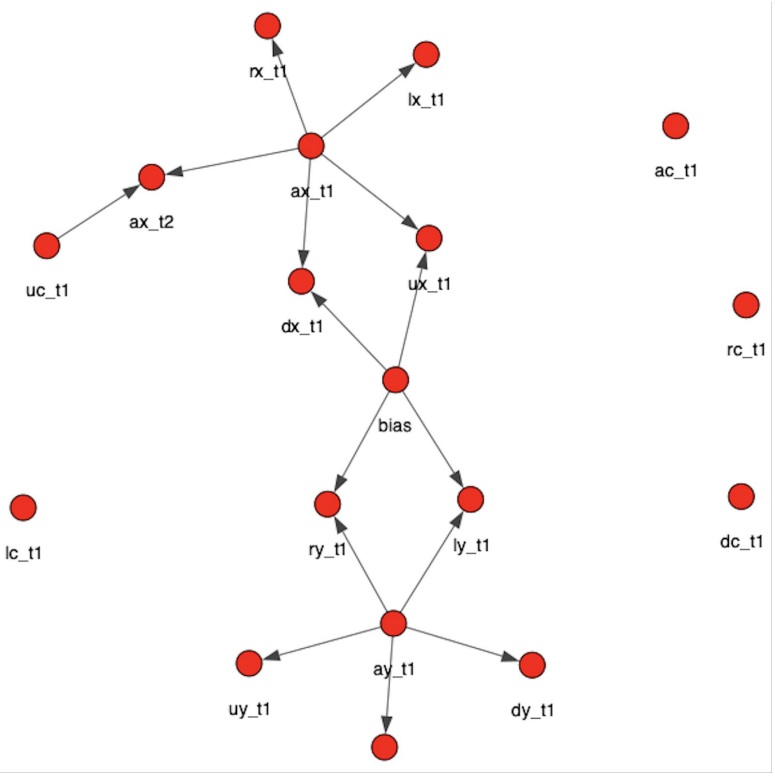
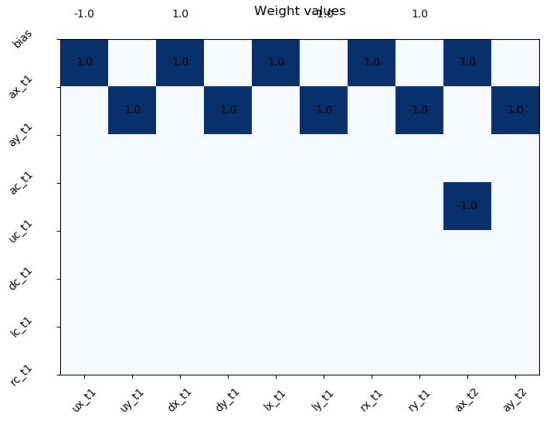
Causal Structural Learning

1. Formulation of generative process of toy game problem as structural equation model.

   (a) Ground Truth Dynamic Bayesian Graph for action = UP



   (b) Ground Truth Weight Parameter Matrix



Structure Learning using Continuous Optimization

DAGs with NO TEARS: Continuous Optimization for Structure Learning: Xun Zheng, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing

1. Conversion of combinatorial highest-scoring structural learning objective to continuous objective function by formulating acyclicity constraint as smooth function.

$$\min_{W \in \mathbb{R}^{d\times d}} F(W) \quad \text{subject to} \quad \mathsf{G}(W) \in \mathsf{DAGs} \iff \min_{W \in \mathbb{R}^{d\times d}} F(W) \quad \text{subject to} \quad h(W) = 0,$$

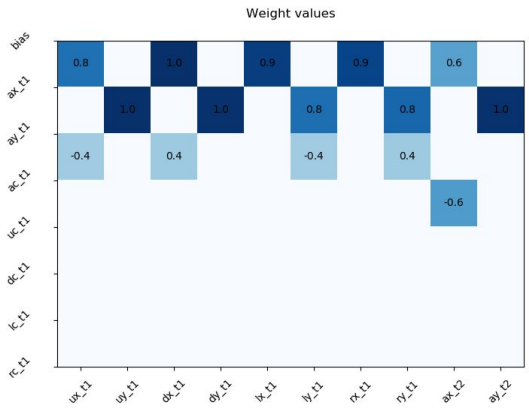
Objective function consists of four parts.

(1) Score function: L2 loss for linear structural equation models
(2) L1 Weight penalty for sparsity
(3) h(W): trace of matrix exponential which enforces acyclicity constraint over matrix W.
(4) Augmented Lagrangian which uses quadratic penalty for h(W)

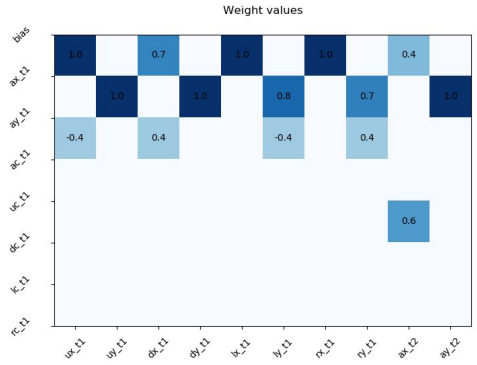
$$\min_{\boldsymbol{w}\in\mathbb{R}^p} f(\boldsymbol{w}) + \lambda\|\boldsymbol{w}\|_1,$$
$$\text{where} \quad f(\boldsymbol{w}) = \ell(W;\mathbf{X}) + \frac{\rho}{2}|h(W)|^2 + \alpha h(W)$$

2.

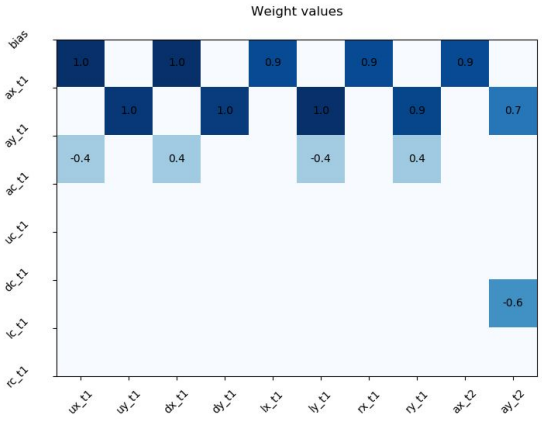
Estimated W using above structure learning algorithm for action up



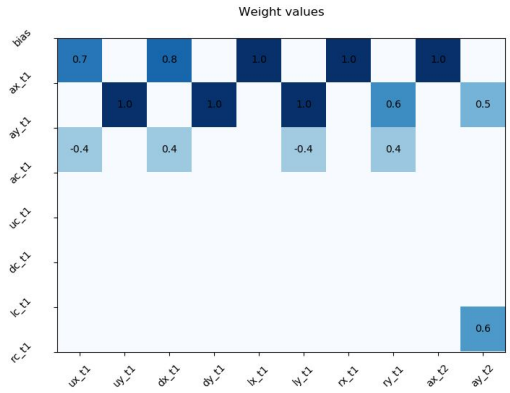
Action down



Action left



Action right



References:

1. https://blog.ml.cmu.edu/2020/04/10/learning-dags-with-continuous-optimization/
2. *DAGs with NO TEARS: Continuous Optimization for Structure Learning: Xun Zheng, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing*
3. Learning Sparse Nonparametric DAGs: Xun Zheng, Chen Dan, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing