A Graphical Criterion for the Identification of Causal Effects in Linear Models

Carlos Brito and Judea Pearl

{fisch,judea}@cs.ucla.edu.

Cognitive Systems Laboratory

University of California, Los Angeles

Introduction

- Find and estimate a model for data obtained from measurements on observable variables.

Introduction

- Find and estimate a model for data obtained from measurements on observable variables.
- Structural Equation Models (SEM)

  Set of linear equations.

Introduction

- Find and estimate a model for data obtained from measurements on observable variables.
- Structural Equation Models (SEM)

  Set of linear equations.
- Natural causal interpretation

History

- Method of Path Coefficients (Wright, 34)
- Economics (Fisher, 66; Bowden & Turkington, 84)
- Social Sciences (Duncan, 75)
- Graphical Methods (Spirtes etal,93; Pearl,95; McDonald,97)

Anatomy and Causal Interpretation of a linear equation

$$Y = \beta X + e$$

Anatomy and Causal Interpretation of a linear equation

$$Y = \beta X + e$$

- Possible (direct) causal influence of X on Y.

Anatomy and Causal Interpretation of a linear equation

$$Y = \beta X + e$$

- Possible (direct) causal influence of X on Y.
- Absence of (direct) causal influence of other variables on Y.

Anatomy and Causal Interpretation of a linear equation

$$Y = \beta X + e$$

- Possible (direct) causal influence of X on Y.
- Absence of (direct) causal influence of other variables on Y.
- β quantifies the Causal-Effect of X on Y.

Anatomy and Causal Interpretation of a linear equation

$$Y = \beta X + e$$

- Possible (direct) causal influence of X on Y.
- Absence of (direct) causal influence of other variables on Y.
- β quantifies the Causal-Effect of X on Y.
- e represents unobserved background factors.

Linear Models

$$\begin{cases} Y_j = \sum_i c_{ji} Y_i + e_j & , j = 1, \ldots, n \\ [\Psi_{ij}] = Cov(e_i, e_j) \end{cases}$$

$$\begin{cases} Y_j = \sum_i c_{ji} Y_i + e_j & , j = 1, \ldots, n \\ [\Psi_{ij}] = Cov(e_i, e_j) & \end{cases}$$

Example:

$$\begin{aligned} Y_1 &= e_1 \\ Y_2 &= e_2 \\ Y_3 &= aY_1 + e_3 \\ Y_4 &= bY_1 + cY_2 + e_4 \\ Y_5 &= dY_3 + fY_4 + e_5 \end{aligned}$$

$$\begin{aligned} Cov(e_1, e_2) &= \alpha \\ Cov(e_2, e_3) &= \beta \\ Cov(e_4, e_5) &= \gamma \end{aligned}$$

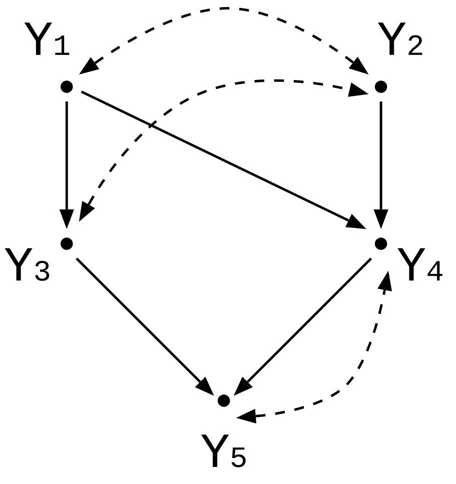
Linear Models

$$\begin{cases} Y_j = \sum_i c_{ji} Y_i + e_j & , j = 1, \ldots, n \\ [\Psi_{ij}] = Cov(e_i, e_j) \end{cases}$$

Example:

$$\begin{aligned} Y_1 &= e_1 \\ Y_2 &= e_2 \\ Y_3 &= aY_1 + e_3 \\ Y_4 &= bY_1 + cY_2 + e_4 \\ Y_5 &= dY_3 + fY_4 + e_5 \end{aligned}$$

$$\begin{aligned} Cov(e_1, e_2) &= \alpha \\ Cov(e_2, e_3) &= \beta \\ Cov(e_4, e_5) &= \gamma \end{aligned}$$



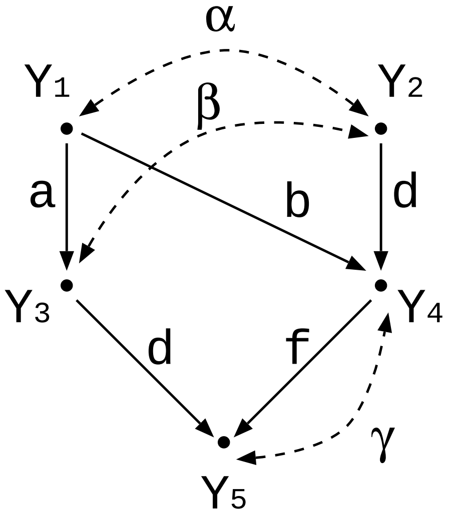
structure

$$\begin{cases} Y_j = \sum_i c_{ji} Y_i + e_j & , j = 1, \ldots, n \\ [\Psi_{ij}] = Cov(e_i, e_j) & \end{cases}$$

Example:

$$\begin{aligned} Y_1 &= e_1 \\ Y_2 &= e_2 \\ Y_3 &= aY_1 + e_3 \\ Y_4 &= bY_1 + cY_2 + e_4 \\ Y_5 &= dY_3 + fY_4 + e_5 \end{aligned}$$

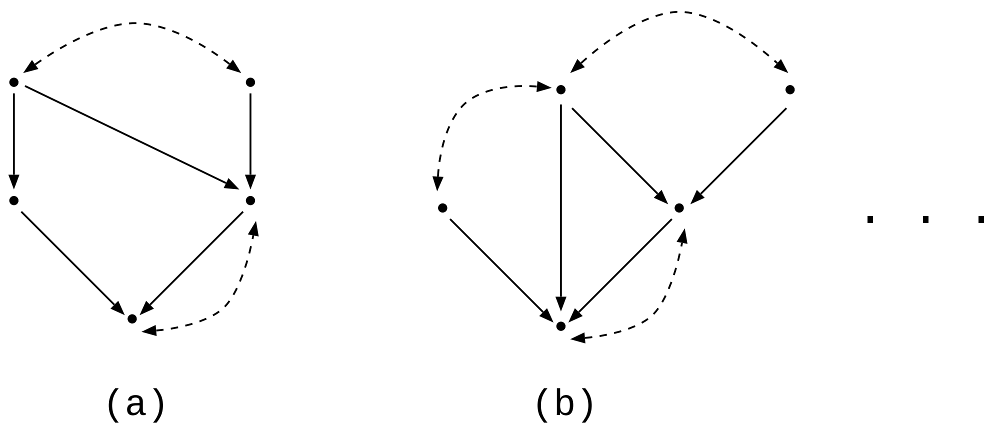
$$\begin{aligned} Cov(e_1, e_2) &= \alpha \\ Cov(e_2, e_3) &= \beta \\ Cov(e_4, e_5) &= \gamma \end{aligned}$$



structure + parameters

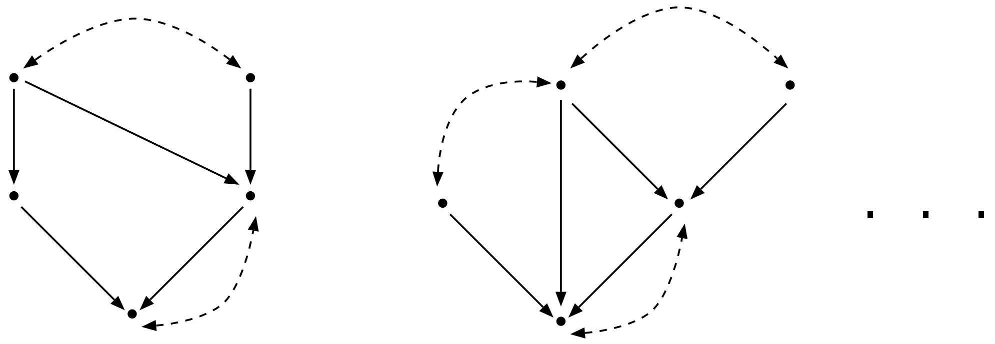
Problems

- Find and Test the model structure



Problems

- Find and Test the model structure



(a) (b)

- The Identification Problem
  - Computing the parameters

The Identification Problem

- Given:
  - Covariance matrix Σ over $\{Y_1, \ldots, Y_n\}$
  - Structure of the model

The Identification Problem

- Given:
  - Covariance matrix Σ over $\{Y_1, \ldots, Y_n\}$
  - Structure of the model
- Want to compute the parameters

The Identification Problem

- Given:
  - Covariance matrix Σ over $\{Y_1, \ldots, Y_n\}$
  - Structure of the model
- Want to compute the parameters
- Existence and Uniqueness

The Identification Problem

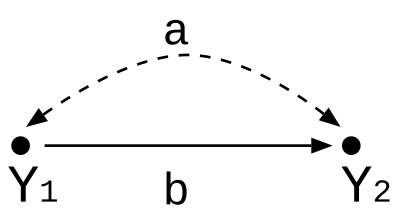
- Given:
  - Covariance matrix Σ over $\{Y_1, \ldots, Y_n\}$
  - Structure of the model
- Want to compute the parameters

- Existence and Uniqueness
  - No parametrization makes the model generate Σ.

The Identification Problem

- Given:
  - Covariance matrix Σ over $\{Y_1, \ldots, Y_n\}$
  - Structure of the model
- Want to compute the parameters

- Existence and Uniqueness
  - No parametrization makes the model generate Σ.
  - There may be more than one solution.



Traditional Approaches

- Based on algebraic manipulations of the equations defining the model

Traditional Approaches

- Based on algebraic manipulations of the equations defining the model
- Rely on conditional independencies implied by the model

Traditional Approaches

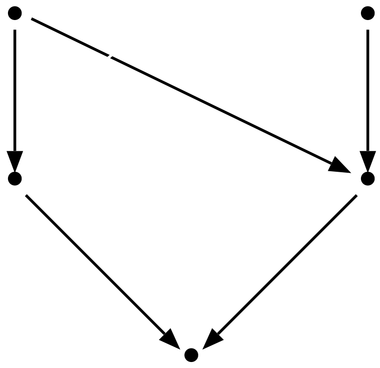
- Based on algebraic manipulations of the equations defining the model
- Rely on conditional independencies implied by the model
- Rank and Order Conditions [Fisher, 66]

Traditional Approaches

- Based on algebraic manipulations of the equations defining the model
- Rely on conditional independencies implied by the model
- Rank and Order Conditions [Fisher, 66]
- Instrumental Variables [Bowden & Turkington, 84]

Graphical Criteria

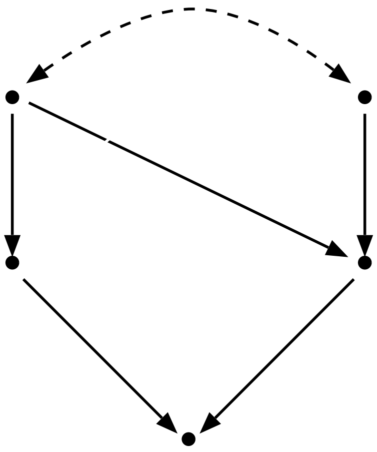
- McDonald's Regressional Hierarchy (McDonald, 1997)



(a) No bidirected edges.

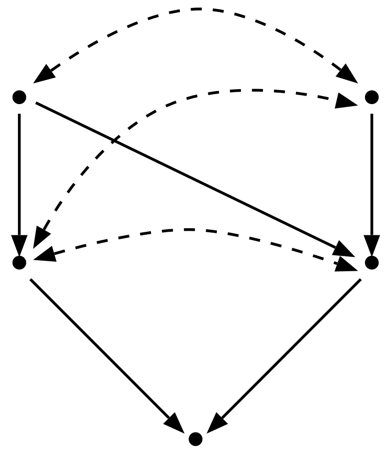
Graphical Criteria

- McDonald's Regressional Hierarchy (McDonald, 1997)



(b) Bidirected edges only between root variables.

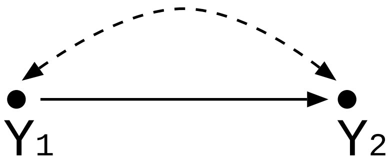
- McDonald's Regressional Hierarchy (McDonald, 1997)



(c) Bidirected edges only between not causally related variables.

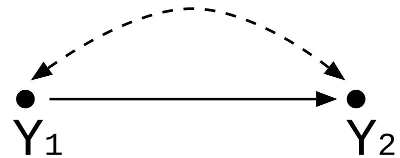
Graphical Criteria (contd)

- Bow-pattern

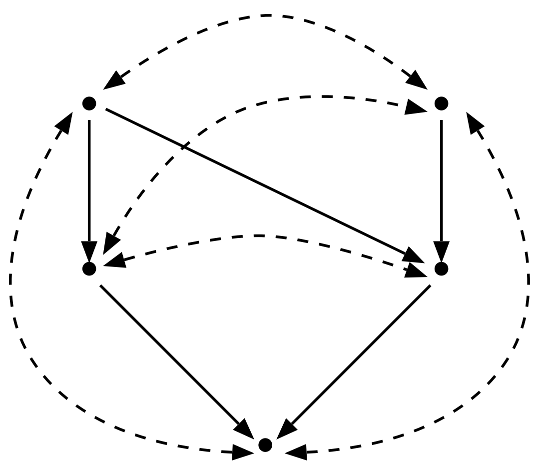


Graphical Criteria (contd)

- Bow-pattern



- Bow-free models (Brito & Pearl, 2002)



Goal

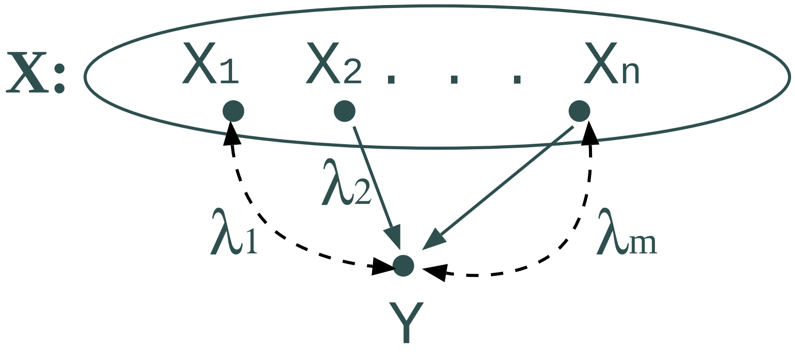
- Find graphical conditions on the structure of a general model that guarantee the identification of all parameters.

Strategy

- Fix a variable Y, and assume that the parameters of edges at depth lower than Y are all identified.

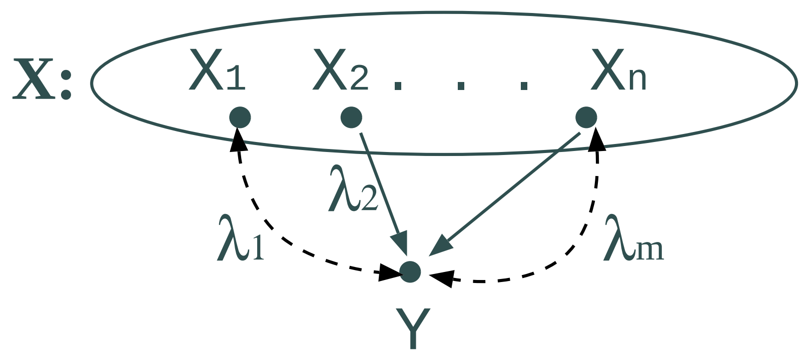
Strategy

- Fix a variable Y, and assume that the parameters of edges at depth lower than Y are all identified.



Strategy

- Fix a variable Y, and assume that the parameters of edges at depth lower than Y are all identified.



- Establish graphical conditions on the causal graph such that the parameters $\lambda_1, \ldots, \lambda_m$ are identified.

Wright's Method of Path Coefficients

- Correlation coefficient of X and Y can be expressed as:

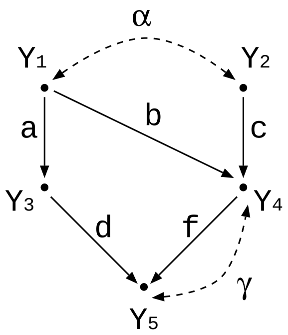
$$\sigma_{X,Y} = \sum_{\text{unblocked paths } p_l} T(p_l)$$

Wright's Method of Path Coefficients

- Correlation coefficient of X and Y can be expressed as:

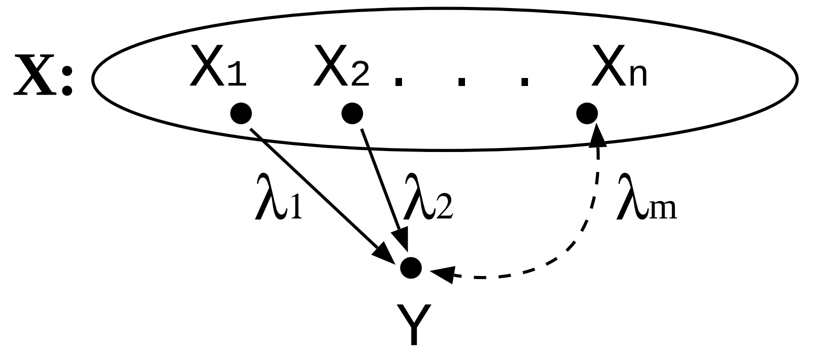
$$\sigma_{X,Y} = \sum_{\text{unblocked paths } p_l} T(p_l)$$

- Example:



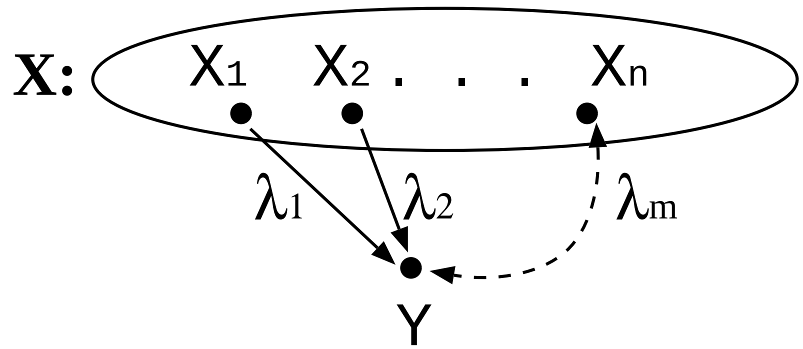
$\rho_{12} = \alpha$
$\rho_{14} = b + \alpha c$
$\rho_{34} = ab + a\alpha c$
$\rho_{15} = ad + bf + \alpha cf$
$\rho_{45} = f + \gamma + c\alpha ad + bad$

Earlier Result (Brito & Pearl, 2002)



Φ = System of Wright's equations for each pair (X_i, Y).

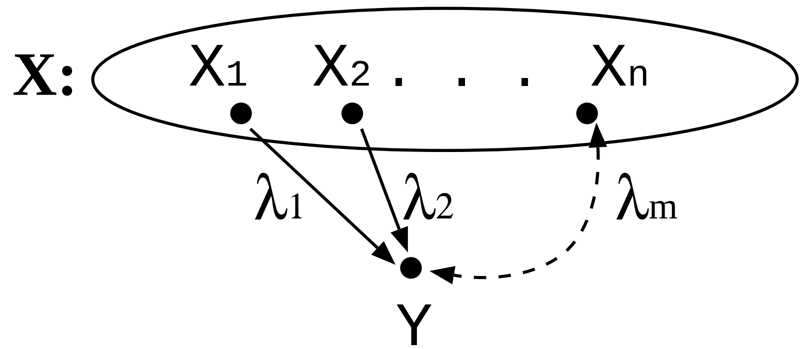
Earlier Result (Brito & Pearl, 2002)



Φ = System of Wright's equations for each pair (X_i, Y).

THEOREM: Equations in Φ are linearly independent w.r.t. λ_i's.

Earlier Result (Brito & Pearl, 2002)

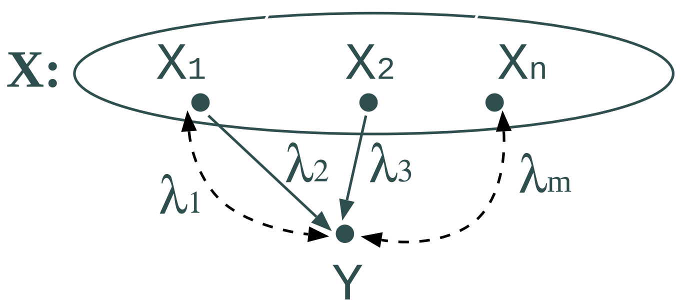


Φ = System of Wright's equations for each pair (X_i, Y).

THEOREM: Equations in Φ are linearly independent w.r.t. λ_i's.

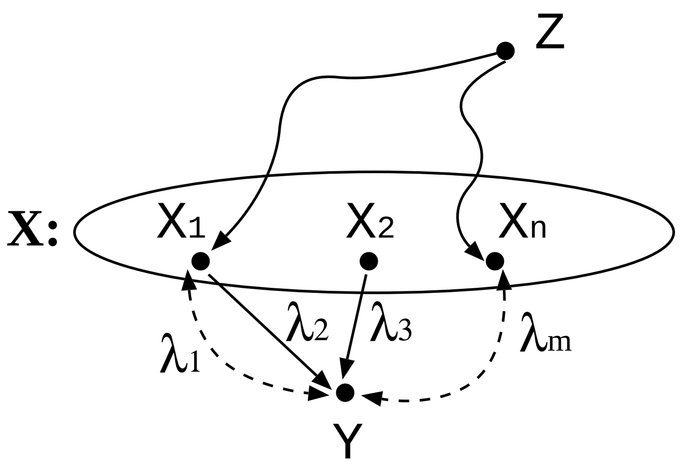
COROLLARY: Bow-free models are completely identifiable.

Auxiliary Variables



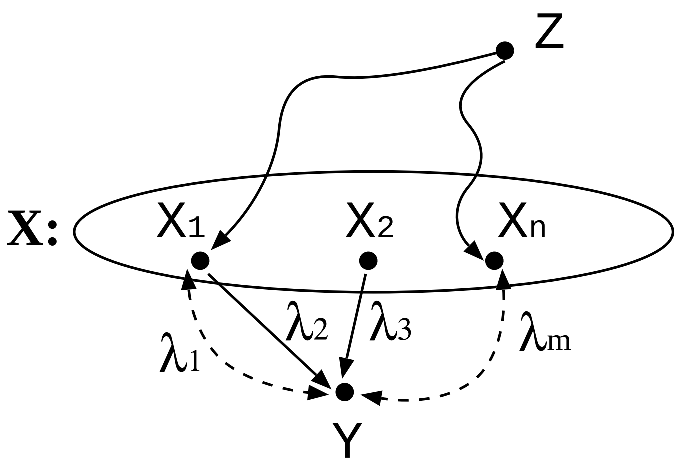
- If Bow-patterns are present, then Φ does not provide enough equations.

Auxiliary Variables



- If Bow-patterns are present, then Φ does not provide enough equations.
- We need additional variables.

Auxiliary Variables



- If Bow-patterns are present, then Φ does not provide enough equations.
- We need additional variables.

DEFINITION: Variable Z is an Auxiliary Variable iff Wright's equation for (Z, Y) is linearly indenpendent from Φ.

Graphical Criterion:

A Chain (or, auxiliary chain) connecting Z and Y is a sequence of paths between $Z, X_1, \ldots, X_k, Y$ such that

(i) there is an unblocked path p_j between X_j and X_{j+1} pointing to both of them;

(ii) there is an unblocked path p_k between X_k and Z pointing to X_k;

(iii) If some $X_l \in \mathbf{X}$ is an intermediate variable of path p_j, then X_l is connected to Y by a bidirected edge and subpath $p_j[X_l \sim X_{j+1}]$ points to X_l.

Graphical Criterion:

A Chain (or, auxiliary chain) connecting Z and Y is a sequence of paths between $Z, X_1, \ldots, X_k, Y$ such that

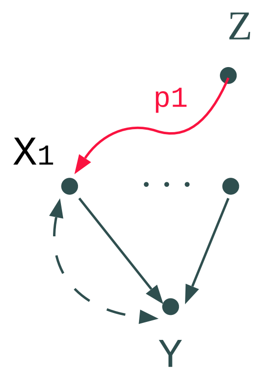
(i) there is an unblocked path p_j between X_j and X_{j+1} pointing to both of them;

(ii) there is an unblocked path p_k between X_k and Z pointing to X_k;

(iii) If some $X_l \in \mathbf{X}$ is an intermediate variable of path p_j, then X_l is connected to Y by a bidirected edge and subpath $p_j[X_l \sim X_{j+1}]$ points to X_l.

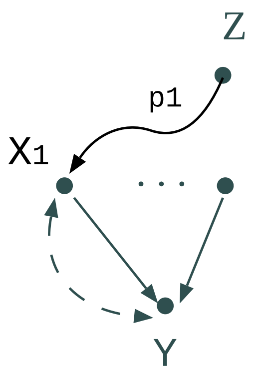
THEOREM: Variable Z is an Auxiliary Variable if and only if there is a chain connecting Z and Y.

Intuition and Examples

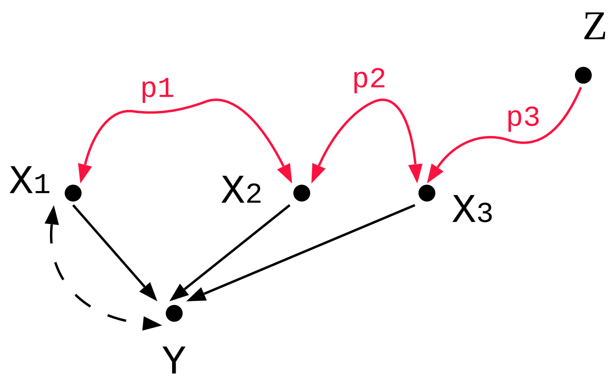


simple case

Intuition and Examples

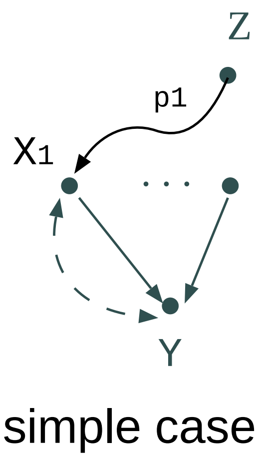


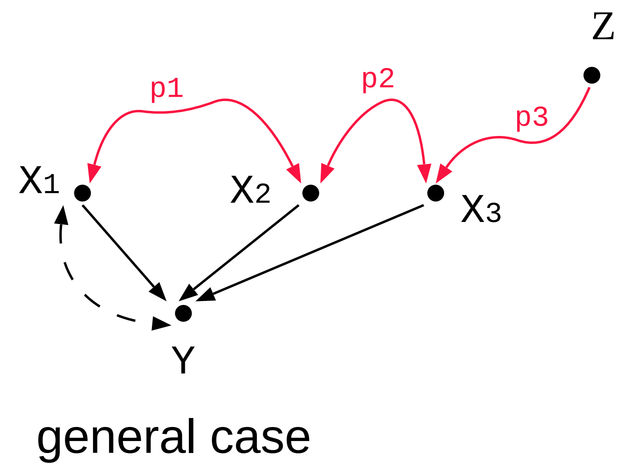
simple case



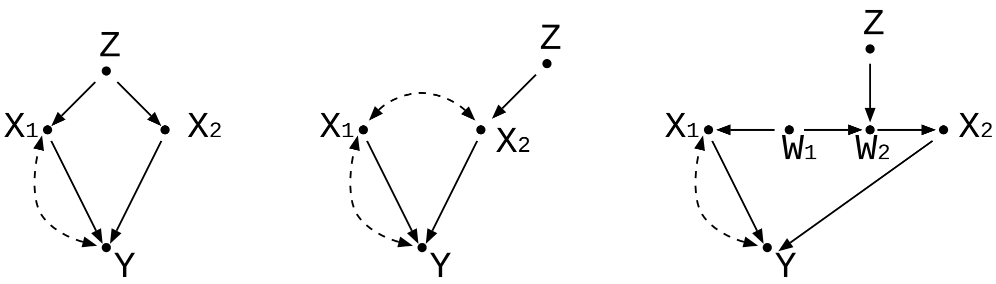
general case

Intuition and Examples





- Examples:



(a) (b) (c)

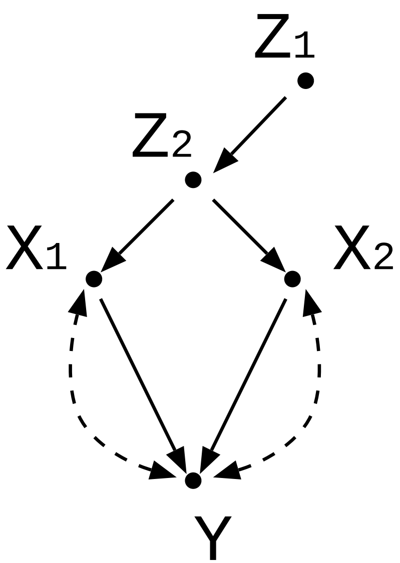
GAV Criterion

- If more than one Bow-pattern is present, then we need more than one Auxiliary Variable.

GAV Criterion

- If more than one Bow-pattern is present, then we need more than one Auxiliary Variable.
- Not every set of Auxiliary variables allow Identification!

- If more than one Bow-pattern is present, then we need more than one Auxiliary Variable.
- Not every set of Auxiliary variables allow Identification!
- Example:



GAV Criterion:

A set $\{Z_1, \ldots, Z_k\}$ of Auxiliary variables satisfies the GAV criterion if:

(i) C_i is an auxiliary chain between Z_i and X_i;

(ii) For any two auxiliary chains C_i and C_j:

(a) C_i and C_j have no common variables other than Z_i or Z_j;

(b) If Z_i appears in path p_l of C_j, then both the subpath $p_i[Z_i \sim X_{l+1}]$ and the last path of C_i point to Z_i.

GAV Criterion:

A set $\{Z_1, \ldots, Z_k\}$ of Auxiliary variables satisfies the GAV criterion if:

(i) C_i is an auxiliary chain between Z_i and X_i;

(ii) For any two auxiliary chains C_i and C_j:

  (a) C_i and C_j have no common variables other than Z_i or Z_j;

  (b) If Z_i appears in path p_l of C_j, then both the subpath $p_i[Z_i \sim X_{l+1}]$ and the last path of C_i point to Z_i.

THEOREM: The GAV criterion is a sufficient condition for a set **Z** of Auxiliary variables to allow identification.

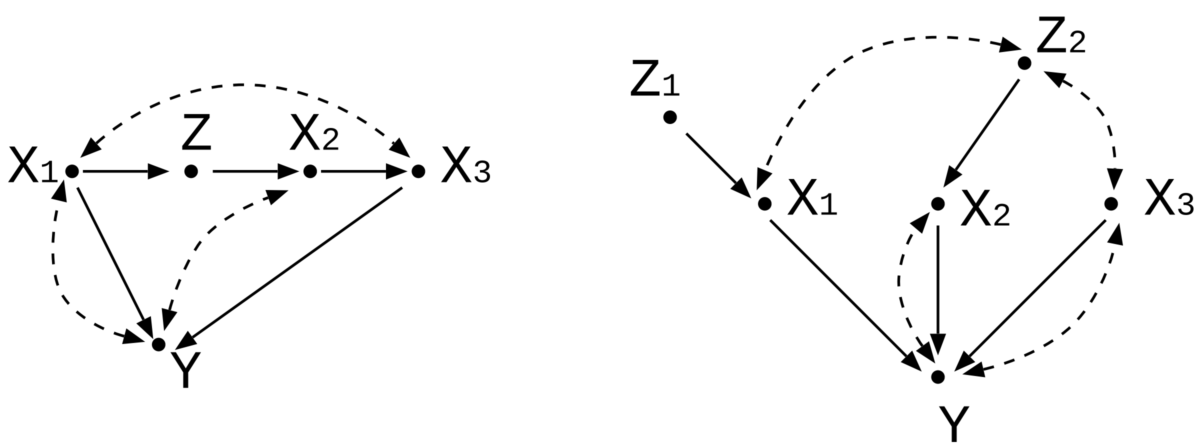
Graphical Criterion for Identification

For each variable Y in the model, check if there is a sufficiently large set of Auxiliary Variables satisfying the GAV criterion.

Graphical Criterion for Identification

For each variable Y in the model, check if there is a sufficiently large set of Auxiliary Variables satisfying the GAV criterion.

Examples:



(a) (b)

Graphical Criterion for Identification

For each variable Y in the model, check if there is a sufficiently large set of Auxiliary Variables satisfying the GAV criterion.

Drawback: The criterion is not simple to be easily verified by inspection of the graph.

Algorithm

- **IDEA:** Reduction to Max-Flow problem.

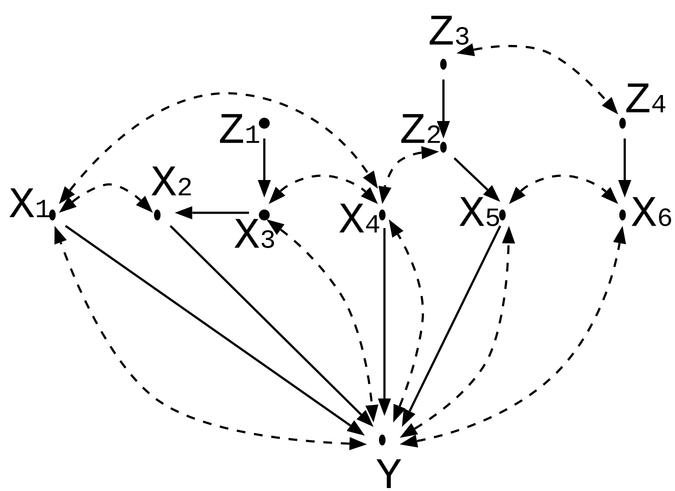
Algorithm

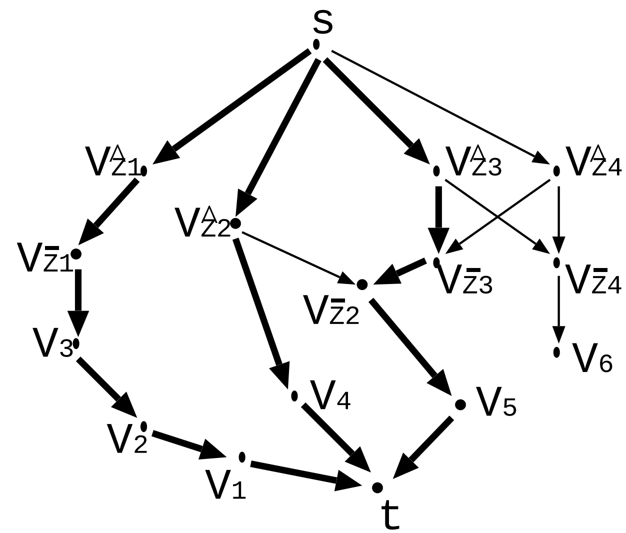
- **IDEA:** Reduction to Max-Flow problem.
- **LEMMA:** If l bow-patterns are present, then we only need to search among variables at distance $\lfloor log(l) \rfloor + 1$ from Y to find Auxiliary variables.

Algorithm

- **IDEA:** Reduction to Max-Flow problem.
- **LEMMA:** If l bow-patterns are present, then we only need to search among variables at distance $\lfloor log(l) \rfloor + 1$ from Y to find Auxiliary variables.
- **THEOREM:** The algorithm is Sound and Complete.

Example


Z3
Z4
Z1
Z2
X1
X2
X3
X4
X5
X6
Y


s
V△Z1
V△Z2
V△Z3
V△Z4
V=Z1
V=Z2
V=Z3
V=Z4
V3
V2
V1
V4
V5
V6
t

Conclusion

- Introduce and completely characterize the concept of Auxiliary variables.

Conclusion

- Introduce and completely characterize the concept of Auxiliary variables.
- A new Graphical criterion for Identification.

Conclusion

- Introduce and completely characterize the concept of Auxiliary variables.
- A new Graphical criterion for Identification.
- The GAV criterion does not make direct use of conditional independence.

Conclusion

- Introduce and completely characterize the concept of Auxiliary variables.
- A new Graphical criterion for Identification.
- The GAV criterion does not make direct use of conditional independence.
- Algorithm to find appropriate Auxiliary variables.