Spiegler (2016, QJE)
Bayesian Networks and Boundedly Rational Expectations

Kyohei Okumura*

September 8, 2018

1 Motivation

- nonrational expectation
- nonrational expectations

2 Approach

- DAG(directed acyclic graph)Causality model
- objective probability distributions $p(x_1, \ldots, x_n)$DAG Rfitsubjective belief $p_R(x_1, \ldots, x_n)$
- (i.e. $p_R(x) \equiv p(x)$)(personal equilibrium)

3 Contribution

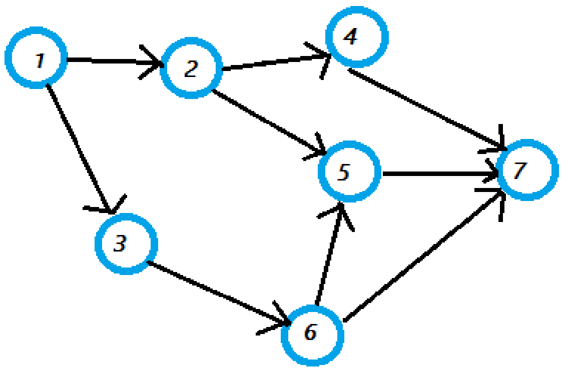
1. Bayesian network factorization formula (bayesian network)
   -
   - $Rp(x_{-1} \mid x_1)p_R(x_{-1} \mid x_1)$
2. Causal/Statistical reasoning
   - reverse causation (): DAG
   - removal of a link (): DAG
   - IllustrationsGeneral Analysis
3. General characterizations of choice behavior
   - rationalirrational
   - causality model Rcausality model R'(:)
4. Bayesian networks as a unifying framework
   - nonrational expectation

*E-mail: kyohei.okumura@gmail.com

4 Model

- $X := X_1 \times \cdots \times X_N$: a finite set of states. $X_1 = A$
- $p \in \Delta(X_1, \ldots, X_n)$: objective probability distribution
- $X_i \widetilde{X}_i$ $(i \in [n])$
- causality model: DAG (N, R), $(NR\text{DAG})$
  - the set of nodes $N := \{1, \ldots, n\}$ $(i \in N\widetilde{X}_i)$
  - the set of edges $R := N \times N$.
  - $(i, j) \in N$node $i$node $j i \to j$
  - $:ij$
  - $R(i) := \{j \in N \mid (j, i) \in R\}$: node i
  - $M \subseteq N x_M := (x_i)_{i \in M}$.

e.g. 4.1 (DAG). $N := \{1, 2, \ldots, 7\}$,
$R := \{(1,2), (1,3), (2,4), (2,5), (3,6), (4,7), (5,7), (6,5), (6,7)\}$
$R(5) = \{2, 6\}$, $x_{R(5)} = (x_2, x_6)$, $Descendants(6) = \{5, 7\}$, $NonDescendants(6) = \{1, 2, 3, 4\}$



- DM$pR()p_R$
- $p = p_R$

$$p_R(x) := \prod_{i=1}^{n} p(x_i \mid x_{R(i)})$$

$$\max_{p(x_1)} \sum_{x_{-1}} p_R(x_{-1} \mid x_1) u(x)$$

e.g. 4.2 (p_R). *Fix $N := \{1, 2, 3\}$ and $p \in \Delta(X_1, X_2, X_3)$. Suppose that DM has his subjective DAG $R : 1 \to 2 \leftarrow 3$. Then, he constructs his subjective belief p_R as follows:*

$$p_R(x_1, x_2, x_3) := p(x_1) p(x_3) p(x_2 \mid x_1, x_3)$$

Lem. 4.1 (p_R is a probability distribution). *For any $p \in \Delta(X)$, the function $p_R : X \to [0, 1]$ is also a probability distribution, i.e., $p_R \in \Delta(X)$.*

Proof. Assume w.l.o.g. that $(1,\dots,n)$ are topologically sorted. [1] Then,

$$\begin{aligned}
\sum_x p_R(x) &= \sum_{x_1}\cdots\sum_{x_n}\prod_{i=1}^{n} p(x_i \mid x_{R(i)}) \\
&= \sum_{x_1}\cdots\sum_{x_{n-1}}\prod_{i\le n-1} p(x_i \mid x_{R(i)}) \underbrace{\sum_{x_n} p(x_n \mid x_{R(n)})}_{=1} \\
&= \sum_{x_1}\cdots\sum_{x_{n-1}}\prod_{i\le n-1} p(x_i \mid x_{R(i)}) \\
&= \cdots \\
&= 1
\end{aligned}$$

$\square$

Def. 4.1 (consistent). *p is consistent with R(, or p factorizes over R)*
$\overset{\Delta}{\iff} p(x) = \prod_{i=1}^{n} p(x_i \mid x_{R(i)}) \iff p = p_R$

- objective probability distribution p is consistent with the true DAG R^*.

Ass. 4.1. *node* 1 *is ancestral in both* R *and* R^*, *i.e.*, $R(1) = R^*(1) = \emptyset$.

Def. 4.2 (Conditional Independence). $V := \{V_1,\dots,V_n\}$*: a set of random variables,* $X, Y, Z \subseteq V$.

$$X \perp Y \mid Z \overset{\Delta}{\iff} [p(Y=y, Z=z) > 0 \implies p(X=x \mid Y=y, Z=z) = p(X=x \mid Z=z)]$$

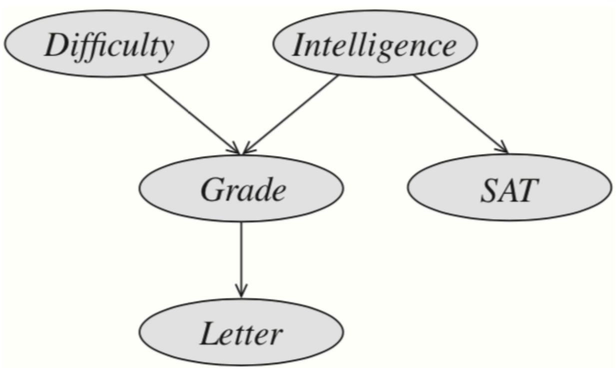
Lem. 4.2 (local independencies). *p factorizes over R iff the following holds:*

$$\widetilde{X}_{NonDescendants(i)} \perp \widetilde{X}_i \mid \widetilde{X}_{R(i)}$$

Cor. 4.1. *Let R be a DAG. Suppose that* $R' \supseteq R$ *and* R' *is also a DAG. If p is consistent with R, then p is also consistent with* R'.

Proof. Suppose that $R' \supseteq R$, and p is consistent with R. Assume w.l.o.g that (N, R) is topologically sorted. Since p is consistent with R, $p(x) = \prod_i p(x_i \mid x_{R(i)})$. Consider the term $p(x_i \mid x_{R(i)})$ for each i. Since R' is a DAG, $x_{R'(i)} = x_{R(i)}$, or $x_{R'(i)} = x_{R(i)} \sqcup x_{N'}$, where $N' \subseteq$ NonDescendants(i); otherwise, R' has a cycle. Then, by Lem.4.2, $p(x_i \mid x_{R'(i)}) = p(x_i \mid x_{R(i)})$. $\square$

e.g. 4.3 (local independencies). *bayesian network structure R (i.e. DAG)pDifficulty: , Intelligence: , Grade: , SAT: SAT, Letter:*
R(Letter) = {Grade}, NonDescendants(Letter) = {Difficulty, Intelligence, SAT}. p is consistent with R [2] *p Difficulty* $\perp$ *Letter* | *Grade*



[1] DAGnode$i \to j \implies i < j$ (i.e. $f : N \to N (i,j) \in R \implies f(i) < f(j)$)$ij > ij \notin R(i)$.
[2] p factorizes over (N, R)DAG$((N, R), p)$*bayesian network*

- historical database interpretation
  - DMDMs
  - true distribution p
  - DMcausal model$i p_R(x_i \mid x_{R(i)})$
  - $p_R(x)(p(a))_a p_R \equiv p()$
  - $p = p_R$causality model Rdata(objective distrib.)
-
  - $p_R(y \mid a)(p(a))_a p_R(y \mid a)$given$(p(a))_a$
  - "trembling": $a p_R(y \mid a) p(a) > 0$

Def. 4.3 (ε-perturbed personal equilibrium). *Fix R and $\varepsilon > 0$. A distribution $p \in \Delta(X)$ with full support on A is an ε-perturbed personal equilibrium*
$\overset{\Delta}{\iff}$

$$\forall a \in A;\ p(a) > \varepsilon \implies a \in \operatorname*{argmax}_{a'} \sum_y p_R(y \mid a') u(a', y)$$

Def. 4.4 (personal eqm.). *$p^* \in \Delta(X)$ is a personal eqm.*
$\overset{\Delta}{\iff}$

$$\exists(\varepsilon_k)_k\ \exists(p_k)_k;\ \varepsilon_k \to 0,\ p_k : \varepsilon_k\text{-perturbed personal equilibrium}, p_k \to p^*$$

Prop. 4.1 (Proposition 2). *For any DAG R, there exists a personal equilibrium.*

Proof. We show the following statement:

$$\forall (p(y \mid a))_{y,a} \exists (p(a))_a;\ p \text{ is PE, where } p(a,y) := p(y \mid a) p(a)$$

Fix $(p(y \mid a))_{y,a}$. Define $Q^\varepsilon \subseteq \Delta(A)$ as follows:

$$Q^\varepsilon := \{\pi \in \Delta(A) \subseteq R^{|A|} \mid \forall a \in A; \pi(a) \geq \varepsilon\}$$

For each $\pi \in Q^\varepsilon$, define p^π, $p^\pi_R(a,y)$ as

$$p^\pi(a,y) := \pi(a) p(y \mid a),\ p^\pi_R(a,y) := \prod_{i=1}^n p^\pi(x_i \mid x_{R(i)})$$

Next, define a correspondence $\mathrm{BR} : Q^\varepsilon \rightrightarrows Q^\varepsilon$ as follows:

$$\mathrm{BR}(\pi) := \operatorname*{argmax}_{\rho \in Q^\varepsilon} \underbrace{\sum_a \rho(a) \sum_y p^\pi_R(y \mid a) u(a,y)}_{=:h(\rho,\pi)}.$$

Lem. 4.3 (Kakutani's theorem). *Suppose the following conditions:*

- *$F : X \rightrightarrows X$ is convex-valued, nonempty-valued and has a closed graph.*
- *X is convex, compact, nonempty.*

Then, there exists $x \in X$ such that $x \in F(x)$.

Lem. 4.4 (Berge's theorem).
- *$f : X \times \Theta \to \mathbb{R}$: continuous.*
- *$\Gamma : \Theta \rightrightarrows X$: compact-valued, continuous.*
- *$v(\theta) := \max_{x \in \Gamma(\theta)} f(x,\theta)$*
- *$x^*(\theta) := \operatorname{argmax}_{x \in \Gamma(\theta)} f(x,\theta)$*

Then, v is continuous, and x^* is u.s.c.

Lem. 4.5 (Sufficient condition for the closed graph). *$F : X \rightrightarrows X$ has a closed graph if F is closed-valued and F is u.s.c.*

Step 1: BR has a fixed point. For sufficiently small $\varepsilon > 0$, Q^ε is convex, compact, and nonempty. $h(\rho) \equiv h(\rho, \pi)$ is linear in ρ; hence, ρ is continuous and quasi-concave in ρ.

- Since h is continuous in ρ and Q^ε is compact, $\mathrm{BR}(\pi) \neq \emptyset$ for all $\pi \in Q^\varepsilon$.
- Since h is continuous, $\mathrm{BR}(\pi)$ is closed.
- Since h is quasi-concave, $\mathrm{BR}(\pi)$ is convex.

Then, we need to show that $\mathrm{BR}(\pi)$ has a closed graph. Since $\mathrm{BR}(\pi)$ is closed-valued, it is sufficient to show that $\mathrm{BR}(\pi)$ is u.s.c. Let $X \times \Theta := Q^\varepsilon \times Q^\varepsilon$ is the statement of Berge's theorem. Since $\Gamma(\theta) \equiv Q^\varepsilon$ (constant), Γ is continuous and compact. We can show that $h(\rho, \pi)$ is continuous not only in ρ but also in pi. ($\because$ $p^\pi(a, y)$ is continuous in π, and then $p^\pi_R(a, y)$ and $p^\pi(y \mid a)$ are also continuous in π.) As h is a function defined on a finite dimensional Euclidean space, h is continuous in (ρ, π). By Berge's theorem, $\mathrm{BR}(\pi)$ is u.s.c. in π; therefore, BR has a fixed point, i.e.,

$$\exists \pi \in Q^\varepsilon;\ \pi \in \mathrm{BR}(\pi).$$

Step 2: p^π is ε-PE. Note that

$$\pi \in \operatorname*{argmax}_{\rho \in Q^\varepsilon} \sum_a \rho(a) \sum_y p^\pi_R(y \mid a) u(a, y).$$

Consider the slightly modified version of the definition of ε-PE:

Def. 4.5 (ε-PE ($\star$)). *$p \in \Delta(X)$ s.t. $\forall a \in A; p(a) \geq \varepsilon$ is ε-PE ($\star$)*
$\overset{\Delta}{\iff}$

$$\forall a \in A;\ p(a) \geq \varepsilon \implies a \in \operatorname*{argmax}_{a'} \sum_y p_R(y \mid a') u(a', y) \tag{1}$$

Lem. 4.6 (The set of PEs remains the same). *Consider two sets of PEs: one is the set of PEs under the original definition of ε-PE, $\mathcal{E}$; the other is the set of PEs under the original definition of ε-PE ($\star$), $\mathcal{E}'$. Then, $\mathcal{E} = \mathcal{E}'$.*

$\mathcal{E}' \subseteq \mathcal{E}$ clearly holds. Fix $p^* \in \mathcal{E}$ and a corresponding sequence $(\varepsilon_k, p_k)_k$. Let $\varepsilon'_k := \min\{\varepsilon_k, p_k(a)\}$. Then, $p'_k \to p^*$ and p'_k is ε'_k-PE. This completes the proof of Lem.4.6.

Here, we show that p^π is a ε-PE ($\star$). Note that π satisfies the condition that $\pi(a) \geq \varepsilon$ for all $a \in A$. Suppose toward contradiction that

$$\exists a \in A; \pi(a) > \varepsilon,\ a \notin \operatorname*{argmax}_{a'} \underbrace{\sum_y p^\pi_R(y \mid a') u(a', y)}_{=:U(a')}$$

Pick some $a^* \in \operatorname{argmax}_{a'} U(a')$. (Since A is finite, we can pick such a^*.) Define $\tilde{\pi} \in Q^\varepsilon$ as follows:

$$\tilde{\pi}(a') = \begin{cases} \pi(a') + \frac{\pi(a) - \varepsilon}{2} & (a' = a^*) \\ \pi(a') - \frac{\pi(a) - \varepsilon}{2} & (a' = a) \\ \pi(a') & \text{o.w.} \end{cases}$$

Note that $\tilde{\pi} \in Q^\varepsilon$ certainly holds. It suffices to check $\tilde{\pi}(a) \geq \varepsilon$:

$$\tilde{\pi}(a) = \frac{2\pi(a) - \pi(a) + \varepsilon}{2} = \frac{\pi(a) + \varepsilon}{2} \geq \varepsilon \quad (\because \pi \in Q^\varepsilon)$$

Observe that $\sum_a \tilde{\pi}(a) U(a) > \sum_a \pi(a) U(a)$. This contradicts $\pi \in \mathrm{BR}(\pi)$. Therefore, p^π is a ε-PE ($\star$).

Step 3: At least one PE p^* exists. So far, we have shown that ε-PE exists (as long as ε is small enough.) Fix some sequence $(\varepsilon^k)_k \subseteq \mathbb{R}$ such that $\varepsilon^k \to 0$. Let p^k be a ε-PE for each k. Note that $(p^k)_k \subseteq \Delta(X) \subseteq \mathbb{R}^{|X|}$. Since $(p^k)_k$ is a sequence in a compact subset of a finite dimensional Euclidean space, $(p^k)_k$ has a convergent subsequence $(p^{k_m})_m$ such that $(p^{k_m})_m \to p^* \in \Delta(X)$. This p^* is PE. $\square$

5 Illustrations

- Reverse causation: Dieter's dilemma
- Coarseness I: Demand for Education
- Coarseness II: Public Policy

5.1 Reverse causation: Dieter's dilemma

- Three variables: a, h, c:
  - DM's choice(diet or not), health outcome(good or bad), chemical level(high or low)
- DM$c, h$

5.1.1 Rational DM

- True DAG: $R^* : a \to c \leftarrow h$
  - $p p(a, h, c) = p(a) p(h) p(c \mid a, h)$
  - DMrational(i.e. causality)

$$\max_a \sum_h \sum_c p(h) p(c \mid a, h) u(a, h, c)$$

5.1.2 Irrational DM

- DMcausality model$R : a \to c \to h$
- $p$personal eqm.$p(a') > 0 a'$

$$a' \in \operatorname*{argmax}_a \sum_h \sum_c p(h \mid c) p(c \mid a) u(a, h, c)$$

5.1.3 Solving for the personal eqm.

- Rpersonal eqm.
-
  - $a, c, h \in \{0, 1\}$
  - $u(a, h, c) = u(a, h) := h - \kappa a$
  - $p(h = 1) = p(h = 0) = 1/2, h \perp a, c = (1 - h)(1 - a)$
- DMrational$p_{R^*}(h \mid a) = p(h) a^* := 0$

Prop. 5.1 (personal eqm. in Dieter's dillemma)**.** *In this case, there is a unique personal eqm p:*

$$p(a = 0) = \begin{cases} 0 & (\kappa \leq 1/4) \\ 2 - \frac{1}{2\kappa} & (\kappa \in (1/4, 1/2)) \\ 1 & (\kappa \geq 1/2) \end{cases}$$

Proof. personal eqm. $p\beta := p(a=0) \in [0,1]p$specification

$$p(c=0 \mid a=1) = 1, p(c=0 \mid a=0) = \frac{1}{2}, p(h=1 \mid c=1) = 0, p(h=1 \mid c=0) = \frac{1}{2-\beta}$$

$$\begin{aligned} p_R(h=1 \mid a=0) &= p(h=1 \mid c=0)p(c=0 \mid a=0) + p(h=1 \mid c=1)p(c=1 \mid a=0) \\ &= \frac{1}{2-\beta}\frac{1}{2} \end{aligned}$$

$$\begin{aligned} p_R(h=1 \mid a=1) &= p(h=1 \mid c=0)p(c=0 \mid a=1) + p(h=1 \mid c=1)p(c=1 \mid a=1) \\ &= \frac{1}{2-\beta} \end{aligned}$$

$\sum_h p(h \mid a)u(a,h)a$

$$\begin{aligned} \sum_h p_R(h \mid a'=0)u(a'=0,h) &= p_R(h=1 \mid a'=0)\cdot 1 \\ &= \frac{1}{2}\frac{1}{2-\beta} \qquad \text{(E0)} \\ \sum_h p_R(h \mid a'=1)u(a'=1,h) &= \frac{1}{2-\beta}(1-\kappa) + \left(1 - \frac{1}{2-\beta}\right) \\ &= \frac{1}{2-\beta} - \kappa \qquad \text{(E1)} \end{aligned}$$

Case (i): $\beta \in (0,1)$ $\beta > \varepsilon, 1-\beta > \varepsilon\varepsilon > 0$personal eqm.(E0) = (E1)

$$\therefore \quad \beta = 2 - \frac{1}{2\kappa}$$

personal eqm.$\varepsilon_k \to 0 p_k := (\beta, 1-\beta)kp_k\varepsilon_k$-perturbed personal eqm$p_k \to p$ok

Case (ii): $\beta = 0$ $1-\beta > \varepsilon$ ε(E0) $\leq$ (E1) (E0) $\leq$ (E1) $\iff \kappa \leq 1/4$. personal eqm.$\kappa \leq 1/4\varepsilon_k \to 0p_k := (0,1)kp_k\varepsilon_k$-perturbed personal eqm$p_k \to p$ok

Case (iii): $\beta = 1$ Case (ii) □

Interpretation:

- dietirrational DMdiet
- DM$a = 0$DM$c, h$negative correlation
- $a \to c \to ha \uparrow \to c \downarrow \to h \uparrow$
- $p(a=1) > 0$
- $a = 1c, h(p(h=1 \mid c=0) = \frac{1}{2-\beta})$

5.2 Coarseness I: Demand for Education

- a, θ, s, w: parent's investment, child's innate ability, school performance, wage
- true DAG R^* :

$$\begin{tikzcd} a \arrow[r] & s \arrow[dr] & \theta \arrow[l] \arrow[d] \\ & & w \end{tikzcd}$$

$$\max_a \sum_\theta p(\theta) \sum_s p(s \mid a, \theta) \sum_w p(w \mid \theta, s) u(a, w)$$

- DM's subjective DAG R :

$$\begin{tikzcd} a \arrow[r] & s \arrow[dr] & \theta \\ & & w \end{tikzcd}$$

$$\max_a \sum_s p(s \mid a) \sum_w p(w \mid s) u(a, w)$$

- θ
- $a \in [0, 1], s, \theta, w \in \{1, 0\}$
- $u(a, w) := w - \kappa(a)$
- κ: twice-differentiable, increasing, weakly convex. (i.e. $\kappa' > 0, \kappa'' \leq 0$), $\kappa'(0) = 0, \kappa'(1) \geq 1$.
- $p(s = 1 \mid a, \theta) = a\theta$, $p(w = 1 \mid s, \theta) = \theta\beta_s$ ($\beta_1 > \beta_0$), $p(\theta = 1) = \delta > 0$.

5.2.1 rational DM's choice

$$\max_a \{\delta[a\beta_1 + (1 - a)\beta_0] - \kappa(a)\}$$

- $\kappa'(a^*) = \delta(\beta_1 - \beta_0)a^*$optimal.

5.2.2 irrational DM's choice

Prop. 5.2. *In this case, the parent assigns probability one to some action a^{**} such that*

$$\kappa'(a^{**}) = \delta \left[\delta\beta_1 - \beta_0 \cdot \frac{\delta(1 - a^{**})}{\delta(1 - a^{**}) + 1 - \delta} \right]$$

If κ' is either weakly convex or weakly concave, then a^{**} is unique.
Note that since $\kappa'(a^{**}) < \kappa'(a^*)$, we have $a^{**} > a^*$: the parent overinvests in personal eqm.

Interpretation:

- The parent overinvests because he overly estimates the positive correlation b/w a and w:
  - DMswpure causal effectθ
  - θw()

  $\rightarrow$
- the perceived marginal benefit of investment $\kappa'(a^{**})$eqm. investment a^{**}

- DM$w \perp_R a \mid s$
- perceived causal effect of s on wa
- i.e. true DAGconsistent$pp(w \mid s, a) \neq p(w \mid s)$
- $s = 0a = 1\theta = 0$
- $\mathbb{E}[w \mid s = 1] - \mathbb{E}[w \mid s = 0]$ increases in long-run investment. ($a$given$s = 0\theta = 0\mathbb{E}[w \mid s = 0]a$)
- true distributionpersonal eqm.true DAG, subjective DAG

Proof of Prop.5.2.

$$
\begin{aligned}
\sum_s p(s \mid a) \sum_w p(w \mid s) u(a, w) &= \sum_s p(s \mid a) p(w = 1 \mid s) - \kappa(a) \\
p(s = 1 \mid a) &= \sum_\theta p(\theta) p(s = 1 \mid a, \theta) = \delta a \\
p(s = 1 \mid a) &= \sum_\theta p(\theta) p(s = 1 \mid a, \theta) = \delta a \\
p(w = 1 \mid s = 1) &= \delta \beta_1 \\
p(w = 1 \mid s = 0) &= \frac{p(w = 1, s = 0)}{p(s = 0)}
\end{aligned}
$$

$$
\begin{aligned}
p(w = 1, s = 0) &= \sum_\theta \sum_a p(w = 1, s = 0, a, \theta) \\
&= \sum_\theta \int_a p(\theta) p(w = 1 \mid s = 0, \theta) p(s = 0 \mid \theta, a) d\mu(a) \\
&= (1 - \delta) \int_a \underbrace{p(w = 1 \mid s = 0, \theta = 0)}_{0} p(s = 0 \mid \theta = 0, a) d\mu(a) \\
&+ \delta \int_a \underbrace{p(w = 1 \mid s = 0, \theta = 1)}_{(\beta_0)} \underbrace{p(s = 0 \mid \theta = 1, a)}_{(1-a)} d\mu(a) \\
&= \delta \beta_0 \int_a (1 - a) d\mu(a)
\end{aligned}
$$

$$
\begin{aligned}
p(s = 0) &= \sum_\theta \sum_a p(a, s = 0, \theta) \\
&= \sum_\theta \sum_a p(\theta) p(a) p(s = 0 \mid a, \theta) \\
&= (1 - \delta) \int_a \underbrace{p(s = 0 \mid a, \theta = 0)}_{1} d\mu(a) + \delta \int_a \underbrace{p(s = 0 \mid a, \theta = 1)}_{(1-a)} d\mu(a) \\
&= (1 - \delta) + \delta \int_a (1 - a) d\mu(a)
\end{aligned}
$$

Then,

$$
p(w = 1 \mid s = 0) = \underbrace{\frac{\delta \int_a (1 - a) d\mu(a)}{(1 - \delta) + \delta \int_a (1 - a) d\mu(a)}}_{=:\gamma} \beta_0
$$

Note that $\gamma < \delta$. Hence,

$$
\sum_s p(s \mid a) p(w = 1 \mid s) - \kappa(a) = \delta a \cdot \delta \beta_1 + (1 - \delta a) \gamma \beta_0 - \kappa(a)
$$

FOC is $\kappa'(a) = \delta(\delta \beta_1 - \gamma \beta_0)$ ($\in (0, 1)$). $\square$

5.3 Coarseness II: Public Policy

- a, y, e, z: policy, two macro variables, private sector's expectation of y.
- true DAG R^* :

$$\begin{tikzcd} a \arrow[r] \arrow[dr] & y \arrow[r] & z \\ & e \arrow[ur] & \end{tikzcd}$$

- DM's DAG R :

$$\begin{tikzcd} a \arrow[r] & y \arrow[r] & z \\ & e & \end{tikzcd}$$

6 General Analysis

6.1 Consequentialist Rationality

- personal equilibrium
- ()

6.1.1 Preliminaries

Def. 6.1 (skeleton). *Fix a DAG $\mathcal{G} := (N, R)$. The skeleton of $\mathcal{G}$, $\widetilde{\mathcal{G}} := (N, \widetilde{R})$, is an indirected version of $\mathcal{G}$: formally, $\widetilde{R} := \{(i,j) \in N \times N \mid (i,j) \in R, \text{ or } (j,i) \in R\}$. $(i,j) \in \widetilde{R}$ is sometimes denoted by $i\widetilde{R}j$, or $i - j$.*

e.g. 6.1 (skeleton). $R : i \to j \to k$, $\widetilde{R} : i - j - k$.

Def. 6.2 (clique, ancestral clique). *Fix a DAG (N, R). $M \subseteq N$ is a clique in R*

$\overset{\Delta}{\iff}$

$$\forall i, j \in M;\ i \neq j \implies i\widetilde{R}j.$$

A clique M in R is an ancestral clique when $\forall i \in M;\ R(i) \subseteq M$.

e.g. 6.2 (clique).

- $M_1 := \{5,6,7\}$: *clique, but not ancestral clique.*
- $M_2 := \{2,4,5,7\}$: *not clique.*
- $M_3 := \{1,3\}$: *ancestral clique.*

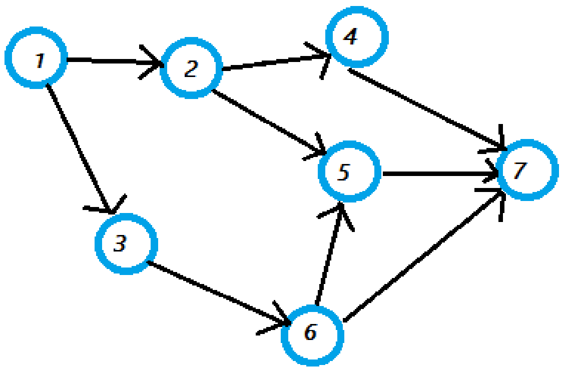


Figure 1: DAG

Def. 6.3 (equivalent). *Fix N. Two DAGs R and Q are equivalent, denoted as $R \sim Q$,*

$\overset{\Delta}{\iff}$

$$\forall p \in \Delta(X); p_R(x) = p_Q(x)$$

We sometimes denote the equivalence class of R as $[R]$.

e.g. 6.3 (equivalent). *$R : 1 \to 2$ and $Q : 2 \to 1$ are equivalent: For any $p \in \Delta(X)$,*

$$p(x_1, x_2) = p(x_2 \mid x_1)p(x_1) = p(x_1 \mid x_2)p(x_2).$$

Def. 6.4 (v-structure). *The v-structure of a DAG R, $v(R)$, is defined as follows:*

$$v(R) := \{(i,j,k) \mid i \to j, j \to k, i \nrightarrow j, j \nrightarrow i\}$$

e.g. 6.4 (v-structure). *Consider the DAG R in Figure 1. $(2,5,6)$ is a v-structure of R; $(5,7,6)$ is not a v-structure in R.*

Prop. 6.1 (Verma and Pearl, 1991). *$R \sim Q \iff [\widetilde{R} = \widetilde{Q}$ and $v(R) = v(Q)]$.*

e.g. 6.5. *$R : 1 \to 2 \to 3$ and $Q : 3 \to 2 \to 1$ are equivalent: $\widetilde{R} = \widetilde{Q} = 1 - 2 - 3$ and $v(R) = v(Q) = \emptyset$. However, $S : 1 \to 2 \leftarrow 3 \nsim R$ because $v(S) = \{(1,2,3)\} \neq \emptyset$.*

6.1.2 Consequentialist Rationality

- $\Delta_R(X) := \{p \in \Delta(X) \mid p \text{ is consistent with } R\}$

Def. 6.5 (Consequentialistically rational). *A DAG R is C-rational w.r.t. true DAG R^** $\overset{\Delta}{\iff}$

$$\forall p, q \in \Delta_{R^*}(X); [\forall x; p(x_{-1} \mid x_1) = q(x_{-1} \mid x_1) \implies \forall x; p_R(x_{-1} \mid x_1) = q_R(x_{-1} \mid x_1)]$$

- R: C-rationaltrue distrib. $p p(x_1) p(x_{-1} \mid x_1) p_R(x_{-1} \mid x_1)$
- $p(x_{-1} \mid x_1)$:given$p(x_1) p(x_{-1} \mid x_1)$

e.g. 6.6 (C-rationality in dieter's dilemma). $p_R(h = 1 \mid a = 0) = \frac{1}{2-\beta} \frac{1}{2}$*dieter's dilemmaRC-rational:* $p p_R(h \mid a) p^*(a) p'(a, h, c) := p(h, c \mid a) p^*(a) \neq p(a, h, c) p'_R(h \mid a) \neq p_R(h \mid a)$

- R^* itself is C-rational w.r.t. R^*.

$\because$) Fix $p, q \in \Delta_{R^*}(X)$ s.t. $p(x_{-1} \mid x_1) = q(x_{-1} \mid x_1)$ for all x. Fix x.
$p_{R^*}(x) = p(x_1) p(x_{-1} \mid x_1)$. $p_{R^*}(x_1) = p(x_1) \sum_{x_{-1}} p(x_{-1} \mid x_1) = p(x_1)$.
Then, $p_{R^*}(x_{-1} \mid x_1) = p(x_{-1} \mid x_1)$. Similarly, $q_{R^*}(x_{-1} \mid x_1) = q(x_{-1} \mid x_1)$. $\square$

- From now on, assume that $R \neq R^*$.

Prop. 6.2 (characterization of C-rationality (Proposition 6)). *R is C-rational w.r.t. R^** $\iff$

$$\forall i > 1;\ 1 \notin R(i) \implies x_i \perp_{R^*} x_1 \mid x_{R(i)}$$

e.g. 6.7 (Dieter's dilemma).
- *True DAG: $R^* : 1 \to 3 \leftarrow 2$*
- *Subjective DAG: $R : 1 \to 2 \to 3$*
- *$i := 3$ $1 \notin R(3), x_3 \not\perp_{R^*} x_1 \mid x_2$.*
- *R is not C-rational w.r.t. R^*.*
- *$R' : 1 \to 3 \quad 2$ (fully coarsed/cursed)*
- *R' is C-rational w.r.t. R^*: $x_2 \perp_{R^*} x_1$.*

- DAG
- d-separation

Proof of Prop.6.2. []

$$
\begin{aligned}
p_R(x_{-1} \mid x_1) = \frac{p_R(x_1, x_{-1})}{p_R(x_1)} &= \frac{p(x_1)\prod_{i\geq 2} p(x_i \mid x_{R(i)})}{\sum_{x'_{-1}} p(x_1)\prod_{i\geq 2} p(x'_i \mid x_{R(i)\cap\{1\}},\ x'_{R(i)-\{1\}})} \\
&= \frac{\prod_{i\geq 2} p(x_i \mid x_{R(i)})}{\sum_{x'_{-1}} \prod_{i\geq 2} p\left(x'_i \mid x_{R(i)\cap\{1\}},\ x'_{R(i)-\{1\}}\right)}
\end{aligned} \tag{2}
$$

$$
\begin{gathered}
p(x_{-1} \mid x_1)\ p(x_1) p_R(x_{-1} \mid x_1) \\
\Longleftrightarrow \forall i > 1;\ 1 \notin R(i) \implies x_i \perp_{R^*} x_1 \mid x_{R(i)}
\end{gathered}
$$

$\Leftarrow)$ (2)$p(x_1) i \geq 2$

$$
p\left(x'_i \mid x_{R(i)\cap\{1\}},\ x'_{R(i)-\{1\}}\right) \tag{$\star$}
$$

$1 \in R(i)$

$$
(\star) = p\left(x'_i \mid x_1, x'_{R(i)}\right)
$$

$(\star) p(x_1)$(???)

$1 \notin R(i) x_i \perp_{R^*} x_1 \mid x_{R(i)}$

$$
\begin{aligned}
(\star) = p\left(x'_i \mid x'_{R(i)}\right) &= \sum_{x''_1} p(x''_1) p(x'_i \mid x''_1, x'_{R(i)}) \\
&= \sum_{x''_1} p(x''_1) p(x'_i \mid x'_{R(i)}) \\
&= p(x'_i \mid x'_{R(i)})
\end{aligned}
$$

$(\star) p(x_1) p(x_1) p_R(x_{-1} \mid x_1)$

$\Rightarrow)$ $i > 1, 1 \notin R(i) i 1 \notin R(i)$

$$
(\star) = p\left(x'_i \mid x'_{R(i)}\right) = \sum_{x''_1} p(x''_1) p(x'_i \mid x''_1, x'_{R(i)})
$$

$x_i \not\perp_{R^*} x_1 \mid x_{R(i)}$ $p(x'_i \mid x''_1, x'_{R(i)})$ x''_1(???) $(\star) p(x''_1)$(?)$p_R(x_{-1} \mid x_1) p(x''_1)$ □

6.2 Behavioral Rationality

- DAG RDMrational – all payoff-relevant variables are causally linked and have no other causes.
- (link)behavioral rationalityviolate

6.2.1 Preliminaries

Def. 6.6 (fully connected). *A directed graph (N, R) is fully connected if $i \to j$ or $j \to i$ holds for all $i, j \in N$.*

Lem. 6.1 (fully connected DAG). *A DAG (N, R) is fully connected $\iff$ R is consistent for all $p \in \Delta(X)$.*

Proof. Assume w.l.o.g that $\{1, 2, \ldots, n\}$ are topologically sorted.

$\Rightarrow$) Fix any x. Then,

$$p(x) = \prod_i p(x_i \mid x_1, \ldots, x_{i-1}) = p_R(x)$$

$\Leftarrow$) We show contraposition. Suppose that R is not fully connected. Then, since R does not have enough its degree of freedom, we can construct p that is not consistent with R. For example, consider $R: 1 \to 2 \to 3$. R is not fully connected because $1 \nrightarrow 3$. Then, we can construct p such that

$$p(x) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_2, x_1) \neq p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_2) = p_R(x)$$

$\square$

Def. 6.7 (d-separation). *Let R be a DAG, and $X, Y, X \subseteq N$.*
A directed path P is d-separated by Z
$\overset{\Delta}{\iff}$

- *P contains a chain $i \to m \to j$ or a fork $i \leftarrow m \to j$ such that $m \in Z$.*
- *P contains an inverted fork $i \to m \leftarrow j$ such that m and the descendants of m are not in Z.*

Z d-separates X and Y $\overset{\Delta}{\iff}$ Z d-separates every path from a node in X to a node in Y. This is denoted by $(X \perp Y \mid Z)_R$. [3]

Prop. 6.3 (Probabilistic Implications of d-Separation). *For any three disjoint subsets of nodes X, Y, Z in a DAG R, and for all probability distributions p,*

1. *If p is consistent with R, then $(X \perp Y \mid Z)_R \implies (X \perp Y \mid Z)_p$*
2. *$(X \not\perp Y \mid Z)_R \implies \exists p;\ (X \not\perp Y \mid Z)_p$.*

6.2.2 Behavioral Rationality

- no restriction on $p \in \Delta(X)$, i.e., assume that true DAG R^* is fully connected.
- Impose some restriction on the set of possible utility functions.

Ass. 6.1 (Restriction on u). *$\exists M \subsetneq N$; $1 \in M$, and u is purely a function of x_M.*

Def. 6.8 (Behaviorally Rational). *A DAG R is B-rational if in every personal eqm. p,*

$$p(x_1) \implies x_1 \in \operatorname*{argmax}_{x_1'} \sum_{x_{-1}} p(x_{-1} \mid x_1) u(x_1', x_{-1})$$

[3]For a probability distribution p, $(X \perp Y \mid Z)_p$ denotes that X and Y are independent conditional on Z.

Prop. 6.4 (Spiegler(2017), Proposition 2). *Let R be a DAG and let $C \subseteq N$.*

$$[\forall p \in \Delta(X)\forall x; p_R(x_C) = p(x_C)] \iff [\exists Q \in [R]; C \text{ is an ancestral clique in } Q].$$

[2018/07/16: $\Leftarrow$ is correct; $\Rightarrow$ is not sure.]

e.g. 6.8. $R : 1 \to 2 \leftarrow 3$. *By Prop.6.1, we can see that* $[R] = \{R\}$. *Since* $\{x_2\}$ *is not an ancestral clique in* R*, by Prop.6.4,* $\exists p \exists x_2; p_R(x_2) \neq p(x_2)$.

Proof of Prop.6.4. See Appendix. □

Prop. 6.5. *The DM is behaviorally rational* $\iff \exists Q \in [R]$; *M is an ancestral clique in Q.*

Proof. [Prop.6.4⇒]
Note that, by assumption, node 1 is an ancestral node in both R and R^*.

$\Leftarrow$) Assume that there exists $Q \in [R]$ such that M is an ancestral clique in Q. By Prop.(6.4), $p_R(x_M) = p(x_M)$. Fix any personal eqm. p. We need to show that p satisfies the following:

$$\forall x_1;\ p(x_1) > 0 \implies x_1 \in \operatorname*{argmax}_{x_1'} \sum_{x_{-1}} p(x_{-1} \mid x_1')u(x).$$

Fix x_1 such that $p(x_1) > 0$. Since u depends only on x_M,

$$\begin{aligned}\operatorname*{argmax}_{x_1'} \sum_{x_{-1}} p(x_{-1} \mid x_1')u(x) &= \operatorname*{argmax}_{x_1'} \sum_{x_{M-\{1\}}} p(x_{M-\{1\}} \mid x_1')u(x_M) \\ &= \operatorname*{argmax}_{x_1'} \sum_{x_{M-\{1\}}} p_R(x_{M-\{1\}} \mid x_1')u(x_M)\ (\because p_R(x_M) = p(x_M)) \\ &= \operatorname*{argmax}_{x_1'} \sum_{x_{-1}} p_R(x_{-1} \mid x_1')u(x)\end{aligned}$$

Since p is personal eqm., $x_1 \in \operatorname{argmax}_{x_1'} \sum_{x_{-1}} p_R(x_{-1} \mid x_1')u(x)$. Therefore, R is B-rational.

$\Rightarrow$) Assume that R is B-rational. By Prop.(6.4), we have $p_R(x_1) = p(x_1)$. Then,

$$p_R(x_{M-\{1\}} \mid x_1) = \frac{p_R(x_M)}{p_R(x_1)} = \frac{p_R(x_M)}{p(x_1)}, \quad p(x_{M-\{1\}} \mid x_1) = \frac{p(x_M)}{p(x_1)}.$$

Hence, $p_R(x_{M-\{1\}} \mid x_1) = p(x_{M-\{1\}} \mid x_1)$ holds if and only if $p_R(x_M) = p(x_M)$ holds.
By Prop.(6.4)[], it is sufficient to show that $p(x_M) \equiv p_R(x_M)$; it suffices to show that $p_R(x_{M-\{1\}} \mid x_1) = p(x_{M-\{1\}} \mid x_1)$. Suppose toward contradiction that $p_R(x_{M-\{1\}} \mid x_1) \neq p(x_{M-\{1\}} \mid x_1)$. Then, we can construct the utility function u under which DM does not choose the optimal action w.r.t. p. (??) □

Interpretation:

- (1) all payoff-relevant variables are causally linked, (2) they have no other causes DMrational
- ((1),(2)pusuboptimal[])
- operationbehavioral rationality

Prop. 6.6 (Proposition 9)**.** *Suppose that R departs from R^*, which is fully connected, by omitting one link $i \to j$. Then,*

$$\textit{DM is B-rational.} \iff j = n,\ i \neq 1.$$

e.g. 6.9.
- $R: 1 \to 3 \leftarrow 2$. $1 \to 2$ *omitted from* R^*. *DM is not B-rational. – double-counting.*
- $R: 1 \to 2 \to 3$. $1 \to 3$ *omitted from* R^*. *DM is not B-rational. – failed to perceive any effect of* x_1
- $R: 2 \leftarrow 1 \to 3$. $2 \to 3$ *omitted from* R^*. *DM is B-rational. – not distinguish direct and indirect effect.*

6.3 Payoff ranking of DAGs

- $\approx$
- DAG – No

e.g. 6.10. • *R: fully connected DAG,* $1 \to 2 \to 3 \to 4$*; u is purely a function of* x_1 *and* x_4.

- R': $2 \to 3$ *removed from R*
- *By Prop.6.6,* R' *is not B-rational:* R' *is weakly dominated by R in terms of expected performance.*
- R'': $2 \to 4$ *removed from* R'.

$$
\begin{tikzcd}
R'': 1 \arrow[r] \arrow[d] \arrow[dr] & 2, \arrow[dl] \\
3 \arrow[r] & 4
\end{tikzcd}
\qquad
\begin{tikzcd}
Q: 1 \arrow[r] \arrow[d] \arrow[dr] & 2 \\
3 \arrow[ur] & 4 \arrow[l]
\end{tikzcd}
$$

- $Q \sim R''$ *(the same skeleton and v-structure).* $\{1,4\}$ *is an ancestral clique in Q.*
- R'' *is B-rational w.r.t.* R^**;* R' *is weakly dominated by* R''.

Ass. 6.2 (For simplicity?). *1 is an isolated node in all relevant true and subjective DAGs.*

Def. 6.9 (Ranking of DAGs). *R is more rational than* R'
$\overset{\Delta}{\iff} \forall p, u, a, a'$;

$$\sum_y p_R(y)u(a,y) > \sum_y p_R(y)u(a',y), \tag{3}$$

$$\sum_y p_{R'}(y)u(a',y) > \sum_y p_{R'}(y)u(a,y) \tag{4}$$

$$\implies \sum_y p(y)u(a,y) > \sum_y p(y)u(a',y) \tag{5}$$

- 2DAG
- R: fully connected, R': not fully connected

Prop. 6.7 (Proposition 10). *Suppose both R and* R' *are not fully connected. Then, neither DAG is more rational than the other.*

Proof. Assume that both R and R' are not fully connected. If $R \sim R'$, the claim holds. Assume $R \nsim R'$.

Suppose toward contradiction that R is more rational than R'. Fix any $p \in \Delta(X)$. Let $q := (p_R(y))_y$ and $r := (p_R(y))_y$. Note that q and r are $k := |Y|$-length probability vectors. Fix any u, a, a'. Let $z^y := u(a,y) - u(a',y)$, $z := (z^y)_y$, and $D := [q \quad -r \quad -p]$. Note that D is a $k \times 3$ matrix. Fix any $\varepsilon > 0$. Let $b := (\varepsilon, \varepsilon, \varepsilon)^\top$.

First, we show the following:

$$\nexists z \in \mathbb{R}^k;\ D^\top z > b \tag{6}$$

Suppose not. Then there exists $z \in \mathbb{R}^k$ such that

$$D^\top z = \begin{bmatrix} q^\top z \\ -r^\top z \\ -p^\top z \end{bmatrix} = \begin{bmatrix} \sum_y p_R(y)(u(a,y) - u(a',y)) \\ -\sum_y p_{R'}(y)(u(a,y) - u(a',y)) \\ -\sum_y p(y)(u(a,y) - u(a',y)) \end{bmatrix} > b$$

This implies

$$\sum_y p_R(y)u(a,y) > \sum_y p_R(y)u(a',y)$$
$$\sum_y p_{R'}(y)u(a',y) > \sum_y p_{R'}(y)u(a,y)$$
$$\sum_y p(y)u(a',y) > \sum_y p(y)u(a,y)$$

This contradicts the assumption that R is more rational than R'. Therefore, (6) must hold.

Next, we apply Gale's theorem:

Lem. 6.2 (Gale's Theorem). *Let $A \in \mathbb{R}^{M\times N}$ and $b \in \mathbb{R}^N$. The following two statements are equivalent:*

1. $\exists x \in \mathbb{R}^M;\ A^\top x \le b$
2. $\forall y \in \mathbb{R}^N;\ y \ge 0, Ay = 0 \implies b^\top y \ge 0$

By (6) and Gale's theorem, we have

$$\exists w \in \mathbb{R}^3;\ w \ge 0, Dw = 0, b^\top w < 0$$

[Spiegler(2016)$w > 0$]

Since $b^\top w < 0$, there exists $j \in \{1,2,3\}$ such that $w_j > 0$. Since $Dw = 0$, for all $i \in [k]$, $w_1 q^i = w_2 r^i + w_3 p^i$, or

$$w_1 p_R(y) = w_2 p_{R'}(y) + w_3 p(y)$$

By summing up w.r.t. i, we have $w_1 = w_2 + w_3$. Hence,

$$w_1 > 0, (w_2 > 0 \text{ or } w_3 > 0)$$

Since $w_1 > 0$, for all y,

$$p_R(y) = \frac{w_2}{w_1} p_{R'}(y) + \frac{w_3}{w_1} p(y)$$

Let $\alpha := w_2/w_1$ and $\beta := w_3/w_1$. Then, by summing up w.r.t. y, we have $\alpha + \beta = 1$. Therefore, we have the following:

$$\forall p\ \exists \alpha \in [0,1]; p_R = \alpha p + (1-\alpha)p_{R'} \tag{7}$$

In case $\alpha < 1$, the proof is done: If p is consistent with R, or $p_R = p$, by (7), we have $p_R = p_{R'}$, and then $p = p_{R'}$; Similarly, if p is consistent with R', then p is also consistent with R: we have the following relationship:

$$p = p_R \iff p = p_{R'}$$

In addition, for any $p \in \Delta(X)$, p_R is consistent with R. Replace p with p_R and apply the procedure to p_R; we have $p_R = \alpha p_R + (1-\alpha)p_{R'}$, and then $p_R = p_{R'}$.

[$\alpha < 1$ for all p, or, $w > 0$ok, fully-connected]

□

7 Variations and Relations to Other Concepts

7.1 Variations

- DAG(Partial cursedness)
- DAG(e.g. Dieters' dilemma)

7.2 Relations to Other Concepts

- Jehiel (2005) Analogy-based expectations
- Esponda (2008) Naive Behavioral Equilibrium
- Eyster and Rabin (2005) Partial cursedness
- Osborne and Rubinstein (1998) S(K) equilibrium

8 Concluding Remarks

8.1 Alternative interpretations of DAG

- Data limitations (cf: Spiegler (2017) Data Monkeys)
- Limited ability to ask the right questions

9 Appendix

Proof of Prop.6.4. [There is an error in the proof in Spiegler(2017).]

If C is empty, the proposition clearly holds; from now on, we assume $C \neq \emptyset$.

First, note that for any DAG R, the following holds:

$$\begin{aligned} p_R(x_C) &= \sum_{x'_{N-C}} p_R(x_C, x'_{N-C}) \\ &= \sum_{x_{N-C}} \prod_{i \in C} p(x_i \mid x_{R(i) \cap C}, x'_{R(i)-C}) \prod_{i \notin C} p(x'_i \mid x_{R(i) \cap C}, x'_{R(i)-C}) \end{aligned} \tag{8}$$

$\Leftarrow$) Fix C such that C is an ancestral clique in some $Q \in [R]$. Note that $R(i) - C = \emptyset$ for all $i \in C$. Then,

$$\prod_{i \in C} p(x_i \mid x_{R(i) \cap C}, x'_{R(i)-C}) = \prod_{i \in C} p(x_i \mid x_{R(i) \cap C}) = p(x_C) \ (\because \text{topological sort})$$

Hence, by (8),

$$p_R(x_C) = p_Q(x_C) = p(x_C) \underbrace{\sum_{x_{N-C}} \prod_{i \notin C} p(x'_i \mid x_{R(i) \cap C}, x'_{R(i)-C})}_{1} = p(x_C).$$

e.g. 9.1. *For example, consider the following DAG:*

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[dr] & 4 \\ & & 3 \end{tikzcd}$$

Let $C := \{1, 2\}$. *Then,*

$$p_R(x_1, x_2) = \sum_{x'_3, x'_4} p_R(x_1, x_2, x'_3, x'_4) = p(x_1, x_2) \sum_{x'_3, x'_4} p(x'_4) p(x'_3 \mid x_2) = p(x_1, x_2)$$

$\Rightarrow$) [We need to make some fix in this direction.]

We show contrapositive: we show the following:

$$[\forall Q \in [R];\ C \text{ is not an ancestral clique in } Q] \implies [\exists p \exists x;\ p_R(x_C) = p(x_C)]$$

Assume that C is not an ancestral clique in any $Q \in [R]$. Fix any $Q \in [R]$. We divide the proof into two cases:

Case (i): In case C is not a clique in Q. In this case, C is not a clique in any $R' \in [R]$. There must be two distinct nodes $i_0, i_1 \in C$ such that $(i_0, i_1) \notin Q$ and $(i_1, i_0) \notin Q$. Consider $p \in \Delta(X)$ such that for every $i \in C \setminus \{i_0, i_1\}$, x_i is independently distributed, whereas x_{i_0} and x_{i_1} are mutually correlated. Then,

$$\prod_{i \in C} p(x_i \mid x_{R(i) \cap C}, x'_{R(i)-C}) = \prod_{i \in C} p(x_i) \ (\because \text{there is no edge b/w } i_0 \text{ and } i_1)$$

$$\prod_{i \notin C} p(x'_i \mid x_{R(i) \cap C}, x'_{R(i)-C}) = \prod_{i \notin C} p(x'_i)$$

$$p_R(x_C) = (8) = \prod_{i \in C} p(x_i) \sum_{i \notin C} \prod_{i \notin C} p(x'_i) = \prod_{i \in C} p(x_i)$$

However,

$$p(x_C) = p(x_{i_0}) p(x_{i_1} \mid x_{i_0}) \prod_{i \in C \setminus \{i_0, i_1\}} p(x_i)$$

Therefore, for some p, $p_R(x_C) \neq p(x_C)$.

Case (ii): C is a clique, but not an ancestral clique in Q. For a DAG R, denote the set of the all v-structures in R as $v(R)$, i.e.,

$$v(R) := \{(i, j, k) \mid i \to j, k \to j, i \nrightarrow k, k \nrightarrow i\}$$

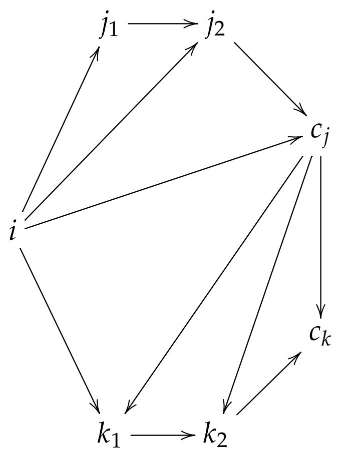
In the original proof, there is a lemma like the following, but the lemma is wrong:

Lem. 9.1. *Let R be a DAG and C be a clique in R. Assume the following two:*

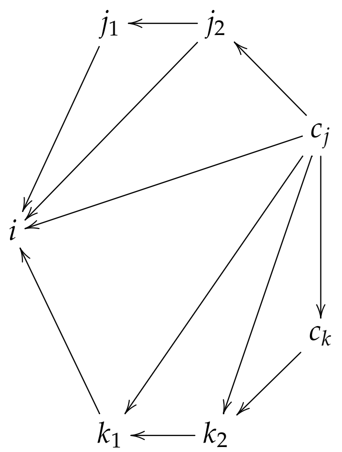
1. *$\forall j \in C$; j has no unmarried parents in R.*
2. *$\forall i \notin C$; if there is a directed path from i to some node $j \in C$ in R, then i has no unmarried parents in R.*

Transform R into another DAG R' by inverting every link along every such path; R and R' has the same v-structure.

e.g. 9.2 (Counter example for Lem.9.1)**.** *Let R be the graph below:*



Let $C := \{c_j, c_k\}$. Note that for all $k \in N \setminus C$ such that k has a path to some $c \in C$, k has no unmarried parents. R' is as follows:



Though $v(R) = \emptyset$, we have $v(R') = \{(j_1, i, k_1), (j_1, i, c_j), (j_2, i, k_1)\}$. Therefore, Lem.9.1 does not hold.

We can consider the modified version of the above lemma:

Lem. 9.2. *Let R be a DAG and C be a clique in R. Assume the following two:*

1. *$\forall j \in C$; j has no unmarried parents in R.*
2. *$\forall i, j \in N$; if there is a directed path from i to some node $c_i \in C$ and a path from j to some node $c_j \in C$ in R, then $i \to j$ or $j \to i$.*

Transform R into another DAG R' by inverting every link along the every path $i \rightsquigarrow c$ such that $i \notin C$ and $c \in C$; then, R and R' has the same v-structure.

For the moment, let us admit Lem.9.2. (I prove it later.)

[I tried to modified the condition in assumption 2 from $\forall i, j \in N$ to $\forall i, j \notin C$, but this does not hold: Below, $(i, j, k) \in v(R)$, but $(i, j, k) \notin v(R')$]

$$R: \begin{tikzcd} i \arrow[dr] & \\ & j \arrow[d] \\ k \in C \arrow[ur] \arrow[r] & c_j \in C \end{tikzcd} \qquad R': \begin{tikzcd} i & \\ & j \arrow[ul] \\ k \in C \arrow[ur] \arrow[r] & c_j \in C \arrow[u] \end{tikzcd}$$

The modified proof for Case (ii) By Lem.9.2, if the two assumptions in Lem.9.2 hold, there should exists $R' \in [R]$ such that C is an ancestral clique in R'; this contradicts the assumption we made at the beginning of the proof.

Hence, one of the following propositions holds:

$$\exists j \in C;\ j \text{ has an unmarried parents in } Q. \tag{P1}$$

$$\exists i, j \in N\ \exists c_i, c_j \in C;\ i \rightsquigarrow_Q c_i, j \rightsquigarrow_Q c_j, i \not\to_Q j, j \not\to_Q i \tag{P2}$$

In case of (P1), the original proof works. From now on, we assume (P1) does not hold and (P2) holds.

First of all, $i \notin C$ or $j \notin C$; otherwise there is an edge between them because C is a clique. Assume w.l.o.g. that $i \notin C$; Q contains the structure as below:

$$\begin{tikzcd} i \arrow[r, squiggly] & c_i \in C \arrow[d, dash] \\ j \arrow[r, squiggly] & c_j \in C \end{tikzcd}$$

Let $P_i \subseteq N$ and $P_j \subseteq N$ are the set of nodes contained in the directed paths from i to c_i and from j to c_j respectively.

Observations:

- $|P_i| \geq 2$. ($\because\ i \notin C$.)
- $|P_j| \geq 1$. (j may be a member of C.)
- c_i and c_j may coincide.
- If $|P_j| = 1$, then $j \neq c_i$; otherwise, $i \to j$.

Consider $p \in \Delta(X)$ and a DAG R^* that satisfy

- p is consistent with R^*.
- $i \notin P_i \cup P_j \implies i$ is an isolated node in R^*.

Consider the subgraph of Q restricted on $P_i \cup P_j$. We name the subgraph Q'.

Case (ii-1): In case $Q'(j) = \emptyset$: Since C is a nonempty clique, $j \notin C$. Since $i \not\leftrightarrow j$, for all $p \in \Delta(X)$, we have $i \not\perp_{p'_Q} j$. Consider $p \in \Delta(X)$ such that $i \perp_p j$. Then, we can apply the same logic in the original proof in this case; we can show the existence of p such that $p(x_C) \neq p_Q(x_C)$ for some x_C.

Case (ii-2): In case $Q'(j) \neq \emptyset$: Fix $k \in Q'(j)$. Since Q' is a DAG, k is not a descendant of j in Q'. We also have $k \neq i$. Since all the nodes in Q' is either the descendant of node i or that of node j, node k is a descendant of node i. Assume w.l.o.g that there is no node along the path from i to k such that the node is a parent of j. (If $|Q'(j)| \geq 2$, then we can take the node k that is closest to i.)

$(i \perp j \mid k)_{Q'}$ **holds:** $\because$) First, take any path $i \rightsquigarrow j$, by the construction of k, k is on that path. Next, we need to check that neither of the following structure is contained in Q':

1. $i \to k \leftarrow j$
2. $$\begin{tikzcd} i \arrow[dr] & & \\ & m \arrow[r, squiggly] & k \\ j \arrow[ur] & & \end{tikzcd}$$

However, since Q' is a DAG and $k \to_{Q'} j$, neither of them holds.

cont. Therefore, there exists $p' \in \Delta(X_{P_i \cup P_j})$ such that $(x_i \not\perp x_j \mid x_k)_{p'}$. Consider the following probability distribution p:

$$p(x) := p'(x_{P_i \cup P_j}) \prod_{l \notin P_i \cup P_j} p(x_l)$$

p_R should satisfy $(x_i \perp x_j \mid x_k)_Q$. This implies

$$\exists p \exists x_C;\ p(x_C) = p_{Q'}(x_C)$$

$\square$

SpieglerLem.9.1

Lem. 9.3. *Let R be a DAG and C be a non-ancestral clique in any $R' \in [R]$. Assume the following two:*

1. *$\forall j \in C$; j has no unmarried parents in R.*
2. *$\forall i \notin C$; if there is a directed path from i to some node $j \in C$ in R, then i has no unmarried parents in R.*

Transform R into another DAG R' by inverting every link along every such path; R and R' has the same v-structure.

Proof of Lem.9.2. We show $v(R) = v(R')$.

Step 1: $v(R) \subseteq v(R')$ Fix any v-structure $(i, j, k) \in v(R)$, $i \to j \leftarrow k$. By assumption 1 in Lem.9.2, we can assume that $j \notin C$. We can also assume that $i \notin C$ or $k \notin C$; otherwise there is an edge between i and k because C is a clique. Assume w.l.o.g that $i \notin C$.

It is sufficient to show that (i, j, k) remains as a v-structure after the inversion. Suppose toward contradiction that (i, j, k) is not a v-structure any more after the inversion. It is necessary that at least one of the edges $i \to j$ and $k \to j$ should be inverted.

Case (1-1): In case $k \notin C$**:** Assume w.l.o.g that $i \to j$ is inverted. Then, there exists some node $c \in C$ such that $i \underset{R}{\rightsquigarrow} c$ [4]; this implies that $i \underset{R}{\rightsquigarrow} c$, and $k \underset{R}{\rightsquigarrow} c$. The graph below summarizes the relationships:

$$\begin{tikzcd}
i \arrow[dr] & & \\
& j \arrow[r, rightsquigarrow] & c \in C \\
k \arrow[ur] & &
\end{tikzcd}$$

However, by assumption 2 in Lem.9.2, there should be an edge between node i and node k; this contradicts the assumption that (i, j, k) is a v-structure in R.

Case (1-2) In case $k \in C$**:** In this case, $k \to j$ is not inverted; then, $i \to j$ should be inverted. Then, by the same logic as in Case (1-1), this leads to a contradiction.

Step 2: $v(R) \supseteq v(R')$ We show that the inversion does not create a new v-structure. Suppose toward contradiction that there exists a triple $(i, j, k) \in v(R) \setminus v(R')$. In this case, the structure as in the below graph should hold in R (c_i and c_k may be the same node.):

$$\begin{tikzcd}
& i \arrow[r, rightsquigarrow] & c_i \in C \\
j \arrow[ur] \arrow[dr] & & \\
& k \arrow[r, rightsquigarrow] & c_k \in C
\end{tikzcd}$$

However, by assumption 2 in Lem.9.2, there should be an edge between node i and node k. A contradiction. □

[4] $i \underset{R}{\rightsquigarrow} j$ denotes that there is a directed path from node i to node j in a DAG R.