

Homework 1
CS 274B: Spring & 2016
Due: April 15, 2016

Problem 1:

I will assume that each random variable can take on d values.

Part A

To satisfy $p(W|X, Y, Z)$, we need $d - 1$ parameters for values of W and each of those values are conditioned on d^3 possible configurations of X, Y, Z , thus we need $d^3(d - 1)$ parameters.

To satisfy $p(Z|X, Y)$, we need $d^2(d - 1)$ parameters because we have $d - 1$ parameter values each conditioned on d^2 configurations.

To satisfy $p(Y|X)$, we need $d(d - 1)$ parameters because we have $d - 1$ parameter values each conditioned on d configurations.

To satisfy $p(X)$, we need $d - 1$ parameters because we have $d - 1$ parameter values.

Our total is thus

$$\frac{d^4 - 1}{d - 1}(d - 1)$$

Simplifying, our final total is

$$d^4 - 1$$

Thus this Bayesian network does not simplify the joint distribution

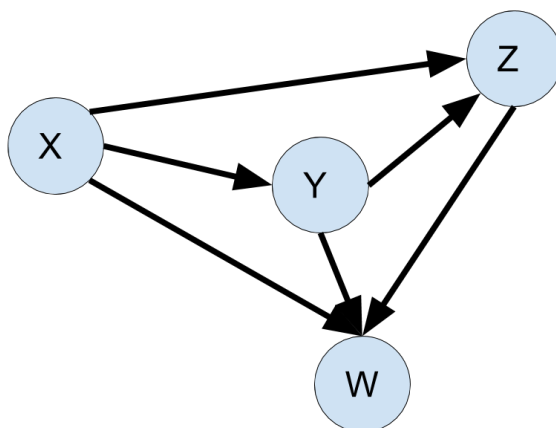


Figure 1: Minimal Directed Graphical Model for Part A

Part B

For each random variable, we need $d - 1$ parameters.

They are all independent

Thus our total is just $4(d - 1)$

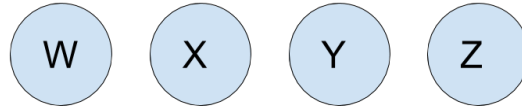


Figure 2: Minimal Directed Graphical Model for Part B

Part C

The variables $p(Z|Y)$, $p(W|Y)$, $p(X|Y)$ each need $d(d - 1)$ parameters since we have $d - 1$ parameter values conditioned on d configurations.

The factor $p(Y)$ needs $d - 1$ parameters

Thus our total is $(3d + 1)(d - 1)$

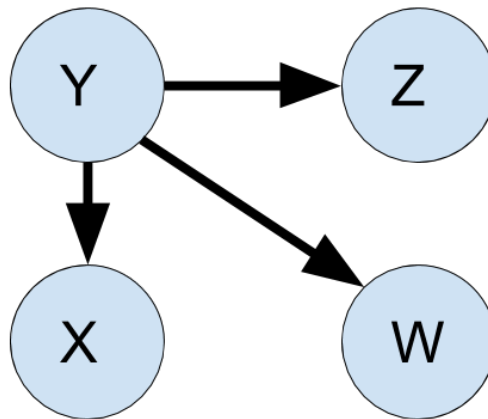


Figure 3: Minimal Directed Graphical Model for Part C

Part D

To satisfy $p(X)$ and $p(Y)$ we need $d - 1$ parameters for each of them

To satisfy $p(W|X)$ we need $d(d - 1)$ parameters since there are $d - 1$ parameter values conditioned on d configurations.

To satisfy $p(Z|X, Y)$ we need $d^2(d - 1)$ parameters since there are $d - 1$ parameter values conditioned on d^2 configurations.

Our total is thus

$$(d^2 + d + 2)(d - 1) = (d^2 + d + 1)(d - 1) + (d - 1) = (d^3 - 1) + (d - 1) = d^3 + d - 2$$

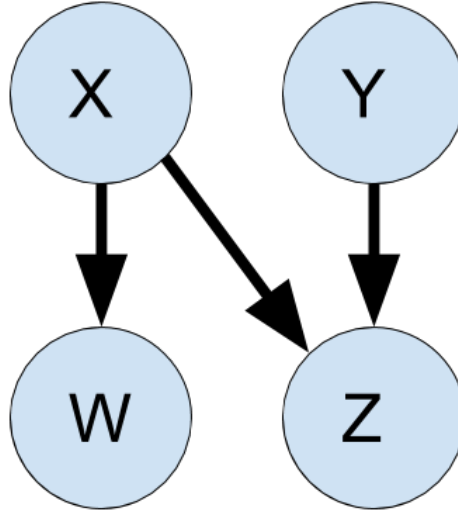


Figure 4: Minimal Directed Graphical Model for Part D

Part E

To satisfy $p(Z)$ we need $d - 1$ parameters

To satisfy $p(Y|Z)$ we need $d(d - 1)$ parameters as there are $d - 1$ values conditioned on d configurations.

To satisfy $p(X|Y)$ we need $d(d - 1)$ parameters for same reason as $p(Y|Z)$

To satisfy $p(W|X)$ we need $d(d - 1)$ parameters for same reason as $p(Y|Z)$

Our total is thus $(3d + 1)(d - 1)$ parameters

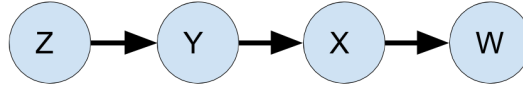


Figure 5: Minimal Directed Graphical Model for Part E

Part F

To satisfy $p(X)$ we need $d - 1$ parameters

To satisfy the other three factors, we need $d(d - 1)$ parameters for each of them for the same reason as $p(Y|Z)$ in part E

Our total is thus $(3d + 1)(d - 1)$ parameters

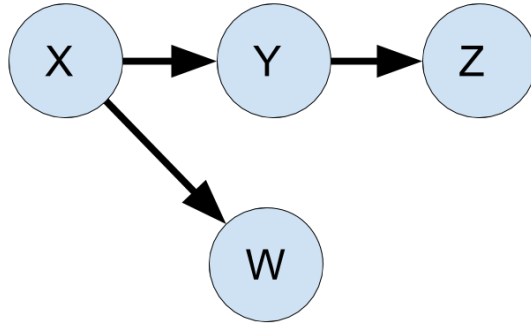


Figure 6: Minimal Directed Graphical Model for Part F

Problem 2:

Part A

No

Knowing the value of *projector_plugged_in* will affect the value of *power_in_wire* and subsequent variables in the network, none of which are *Sam_reading_book*

Further, the value of *Sam_reading_book* is conditioned on the value of *room_light_on*

The variables *room_light_on* and *power_in_wire* are conditionally independent given *power_in_building*

Part B

No

Knowing that *screen_lit_up* is true means that the probability of *ray_says_screen_is_dark* is affected but no other variable ended up affected.

Part C

No

Knowledge of *screen_lit_up* and *projector_plugged_in* affect subsequent variables in the network none of which are *Sam_reading_book*

Part D

If *lamp_works* was observed, then we would update the probabilities for *projector_lamp_on*

We would then have to update the probabilities for *screen_lit_up*

This would cause us to update probabilities for *ray_says_screen_is_dark*

Part E

If we observe just *power_in_projector* then the same variables from Part D will have their probabilities changed.

Problem 3:

Part A

We need to solve the following

$$p(0, 0; \theta) + p(0, 1; \theta) + p(1, 0; \theta) + p(1, 1; \theta) = 1$$

This ends up being the following:

$$\exp(-A(\theta)) + \exp(-A(\theta)) + \exp(\theta_x - A(\theta)) + \exp(\theta_x + \theta_{xy} - A(\theta)) = 1$$

After doing some factoring

$$\frac{\exp(\theta_x) + \exp(\theta_{xy} + \theta_x) + 2}{\exp(A(\theta))} = 1$$

After cross multiplying and solving for $A(\theta)$

$$A(\theta) = \log(\exp(\theta_x) + \exp(\theta_{xy} + \theta_x) + 2)$$

Part B

After letting $\theta_{xy} = 1$ we have the following

$$A(\theta) = \log(\exp(\theta_x) + \exp(1 + \theta_x) + 2)$$

After some factoring

$$A(\theta) = \log(\exp(\theta_x)(1 + \exp(1)) + 2)$$

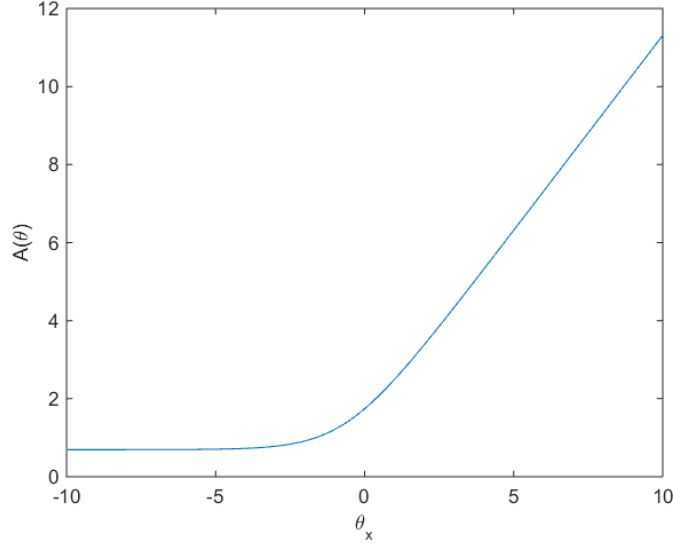


Figure 7: Plot for $A(\theta)$. It appears convex as expected

Part C

This is the partial with respect to θ_x

$$\frac{\partial A}{\partial \theta_x} = \frac{\exp(\theta_x) + \exp(\theta_x + \theta_{xy})}{\exp(\theta_x) + \exp(\theta_x + \theta_{xy}) + 2}$$

This is the partial with respect to θ_{xy}

$$\frac{\partial A}{\partial \theta_{xy}} = \frac{\exp(\theta_x + \theta_{xy})}{\exp(\theta_x) + \exp(\theta_x + \theta_{xy}) + 2}$$

Thus we have

$$\nabla A(\theta) = \left[\frac{\exp(1) + \exp(3)}{\exp(1) + \exp(3) + 2}, \frac{\exp(3)}{\exp(1) + \exp(3) + 2} \right]$$

Approximately

$$\nabla A(\theta) = [0.91937, 0.80978]$$