CSC412 – Assignment #1 Solutions

1 Graphical Model Distributions (8 marks)

	Directed model	Undirected model
1. $\mathbf{x}_4$ is conditionally independent of $\mathbf{x}_5$ given $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$	must be true	could be true
2. $\mathbf{x}_3$ is conditionally *dependent* on $\mathbf{x}_6$ given $\mathbf{x}_4$ and $\mathbf{x}_5$	cannot be true	cannot be true
3. $\mathbf{x}_1$ is marginally independent of $\mathbf{x}_2$	must be true	could be true
4. $\mathbf{x}_4$ is conditionally independent of $\mathbf{x}_5$ given $\mathbf{x}_3, \mathbf{x}_6$	could be true	must be true

- The structure of a graphical model only asserts conditional independence statements about the distribution.
- If the structure does not imply conditional independence between two variables given a set of variables, the two variables could still be conditionally indepedent (for certain parameter settings of the model).

2 Numeric Distributions (3+12 marks)

2.1 Distributions

1. $pA(\mathbf{x}_3, \mathbf{x}_4 | \mathbf{x}_1$ takes its third value)

```
>> z1=squeeze(sum(pA(3,:,:,:),2)); z1=z1/sum(z1(:)); z1

    0.0411    0.0411    0.0411    0.0411    0.0411
    0.0392    0.0392    0.0392    0.0392    0.0392
    0.0381    0.0381    0.0381    0.0381    0.0381
    0.0501    0.0501    0.0501    0.0501    0.0501
    0.0611    0.0014    0.0442    0.0154    0.0356
```

2. $pB(\mathbf{x}_4 | \mathbf{x}_2$ takes its first value)

```
>> z2=squeeze(sum(sum(pB(:,1,:,:),3),1))'; z2=z2/sum(z2(:)); z2

    0.3300    0.1100    0.2200    0.0700    0.2700
```

3. $pC(\mathbf{x}_1)$

```
>> z3=sum(sum(sum(pC,4),3),2)'; z3=z3/sum(z3(:)); z3

    0.2039    0.1863    0.1943    0.2130    0.2026
```

- When computing conditional distributions, please renormalize.
- The command `squeeze` is useful for removing singleton dimensions.

2.2 Conditional independence statements

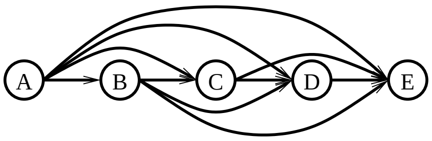
	A	B	C	Code for A
1. $\mathbf{x}_1 \perp\!\!\!\perp \mathbf{x}_2 \,|\, \mathbf{x}_3$	No	Yes	No	`for ii=1:5; rank(sum(pA(:,:,ii,:),4)), end`
2. $\mathbf{x}_1 \perp\!\!\!\perp \mathbf{x}_2 \,|\, \mathbf{x}_3, \mathbf{x}_4$	Yes	Yes	No	`for ii=1:5; for jj=1:5; rank(pA(:,:,ii,jj)), end, end`
3. $\mathbf{x}_1 \perp\!\!\!\perp \mathbf{x}_2$	No	Yes	No	`rank(sum(sum(pA,3),4))`
4. $\mathbf{x}_3 \perp\!\!\!\perp \mathbf{x}_4 \,|\, \mathbf{x}_1, \mathbf{x}_2$	No	Yes	No	`for ii=1:5; for jj=1:5; rank(squeeze(pA(ii,jj,:,:))), end, end`

- If a joint distribution given by a matrix has rank 1, it means it is a product of two vectors, i.e. variables are independent (and vice versa).
- To test whether two variables x, y are conditionally independent given z, make sure $p(x, y|z)$ has rank 1 for all values of z.

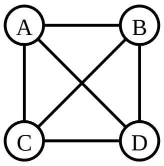
3 Model Conversions (10 marks)

	A	B	C
Topological orderings	(x, y, z)	$(y, x, z), (y, z, x)$	$(y, x, z, w), (y, z, x, w)$
Conditional independencies	$x \perp\!\!\!\perp z \mid y$	$x \perp\!\!\!\perp z \mid y$	$x \perp\!\!\!\perp z \mid y$ and $y \perp\!\!\!\perp w \mid x, z$
Undirected model	$x - y - z$	$x - y - z$	Remove $x \to w$ and $x - y - z - w$ (or remove $z \to w$ instead)

4 Complete Graphs (6 marks)



- No edges can be added because this will create cycles.
- No edges can be removed as it implies a conditional independency which means that distributions not satisfying that conditional independency cannot be represented. If edge $u \to v$ is removed, the conditional independency is $u \perp\!\!\!\perp v \mid$ all parents of u and v.



- No edges can be added as all edges are already in there.
- No edges can be removed as it implies conditional independency which means that distributions not satisfying that conditional independency cannot be represented. If edge $u - v$ is removed, the conditional independecy is $u \perp\!\!\!\perp v \mid$ all other nodes.

5 Incompatible Conditionals (2+0+9 marks; 3 bonus marks)

- The function $f(x, y) = p(x|y)p(y|x)$ is not a valid joint distribution because it does not sum to 1.
- No valid joint distribution $p(x, y)$ exists with the given conditionals. This is a typo on our fault. As a result we won't count any marks for this. However if you can prove that no such distribution exist, you get 3 bonus marks. This is proven using contradiction.
- If $p(x, y) = p(x|y)p(y|x)$, and we assume none of the probabilities are 0, then $p(x, y) = p(x)p(y|x) = p(x|y)p(y|x)$ so $p(x) = p(x|y)$ by cancelling $p(y|x)$ from both sides. This means that x and y must be marginally independent. Now any two distribution with x independent of y will do for the third sub-question.