Branch and Bound for Regular Bayesian Network Structure Learning: Supplementary Material

Joe Suzuki∗ and Jun Kawahara†
∗ Osaka University, Japan. j-suzuki@sigmath.es.osaka-u.ac.jp
† Nara Institute of Science and Technology, Japan. jkawahara@is.naist.ac.jp

Appendix A

Under $c_n(x,y)=c_n(y)$ for each (x,y), we have

$$
\begin{aligned}
& Q^n(X|Y) \geq Q^n(X|YZ) \\
\Longleftrightarrow\quad & \frac{\Gamma(\frac{\alpha\beta}{2})}{\Gamma(n+\frac{\alpha\beta}{2})}\prod_{x,y}\frac{\Gamma(c_n(x,y)+\frac{1}{2})}{\Gamma(\frac{1}{2})}\cdot\frac{\Gamma(\frac{\beta\gamma}{2})}{\Gamma(n+\frac{\beta\gamma}{2})} \\
& \cdot\prod_{y,z}\frac{\Gamma(c_n(y,z)+\frac{1}{2})}{\Gamma(\frac{1}{2})} \\
& \geq \frac{\Gamma(\frac{\alpha\beta\gamma}{2})}{\Gamma(n+\frac{\alpha\beta\gamma}{2})}\prod_{x,y,z}\frac{\Gamma(c_n(x,y,z)+\frac{1}{2})}{\Gamma(\frac{1}{2})} \\
& \cdot\frac{\Gamma(\frac{\beta}{2})}{\Gamma(n+\frac{\beta}{2})}\prod_{y}\frac{\Gamma(c_n(y)+\frac{1}{2})}{\Gamma(\frac{1}{2})} \\
\Longleftrightarrow\quad & \frac{\Gamma(\frac{\alpha\beta}{2})}{\Gamma(n+\frac{\alpha\beta}{2})}\cdot\frac{\Gamma(\frac{\beta\gamma}{2})}{\Gamma(n+\frac{\beta\gamma}{2})} \\
& \geq \frac{\Gamma(\frac{\alpha\beta\gamma}{2})}{\Gamma(n+\frac{\alpha\beta\gamma}{2})}\cdot\frac{\Gamma(\frac{\beta}{2})}{\Gamma(n+\frac{\beta}{2})}\ .
\end{aligned}
$$

The last inequality can be proved using induction w.r.t. n, which completes the proof.

Appendix B

From Stirling's formula $\log\Gamma(z) = z\log z - z + \frac{1}{2}\log\frac{2\pi}{z} + \epsilon(z)$ with $\frac{1}{12z} < \epsilon(z) < \frac{1}{12z+1}$, we have for $\beta=\sigma(S)$

$$
\begin{aligned}
& \log\Gamma(n+\frac{\alpha\beta}{2}) - \log\Gamma(n+\frac{\beta}{2}) \\
= \quad & -(n+\frac{\beta}{2})\log(n+\frac{\beta}{2}) + (n+\frac{\beta}{2}) - \frac{1}{2}\log\frac{2\pi}{n+\frac{\beta}{2}} \\
& +(n+\frac{\alpha\beta}{2})\log(n+\frac{\alpha\beta}{2}) - (n+\frac{\alpha\beta}{2}) \\
& +\frac{1}{2}\log\frac{2\pi}{n+\frac{\alpha\beta}{2}} + O(\frac{1}{n}) \\
= \quad & -n\log\frac{n+\frac{\beta}{2}}{n+\frac{\alpha\beta}{2}} - \frac{\beta}{2}\log(n+\frac{\beta}{2}) \\
& +\frac{\alpha\beta}{2}\log(n+\frac{\alpha\beta}{2}) - \frac{1}{2}\log(n+\frac{\alpha\beta}{2}) \\
& +\frac{1}{2}\log(n+\frac{\beta}{2}) - \frac{\alpha\beta-\beta}{2} + O(\frac{1}{n}) \\
= \quad & \frac{(\alpha-1)\beta}{2}\log n + O(\frac{1}{n})\ ,
\end{aligned}
$$

which means

$$
\begin{aligned}
& -\log\{\prod_{i=1}^{n}\frac{i-1+0.5\beta}{i-1+0.5\alpha\beta}\} \\
= \quad & \log\frac{\Gamma(n+\frac{\alpha\beta}{2})\Gamma(\frac{\beta}{2})}{\Gamma(n+\frac{\beta}{2})\Gamma(\frac{\alpha\beta}{2})} \\
= \quad & \frac{(\alpha-1)\beta}{2}\log n + \log\frac{\Gamma(\frac{\beta}{2})}{\Gamma(\frac{\alpha\beta}{2})} + O(\frac{1}{n})\ .
\end{aligned}
$$

For the second statement, from

$$
\begin{aligned}
& \log\Gamma(c_n(x,s)+\frac{1}{2}) \\
= \quad & c_n(x,s)\log\{c_n(x,s)+\frac{1}{2}\} - \{c_n(x,s)+\frac{1}{2}\} \\
& +\frac{1}{2}\log(2\pi) \\
= \quad & c_n(x,s)\log c_n(x,s) - c_n(x,s) + O(\frac{1}{c_n(x,s)})
\end{aligned}
$$

and

$$
\log\Gamma(c_n(s)+\frac{1}{2}) = c_n(s)\log c_n(s) - c_n(s) + O(\frac{1}{c_n(s)})\ ,
$$

we have

$$
\begin{aligned}
& -\log\{\prod_{i=1}^{n}\frac{c_{i-1}(x_i,s_i)+0.5}{c_{i-1}(s_i)+0.5}\} \\
= \quad & -\log\prod_{s}\{\frac{\prod_x\Gamma(c_n(x,s)+\frac{1}{2})}{\Gamma(c_n(s)+\frac{1}{2})}\}
\end{aligned}
$$

$$
\begin{aligned}
&= -\sum_s \sum_x c_n(x,s) \log \frac{c_n(x,s)}{c_n(s)} + O(1) \\
&= -n \sum_s \frac{c_n(s)}{n} \sum_x \frac{c_n(x,s)}{c_n(s)} \log \frac{c_n(x,s)}{c_n(s)} + O(1) \ .
\end{aligned}
$$

Since $H(X|S) = -\sum_s \frac{c_n(s)}{n} \sum_x \frac{c_n(x,s)}{c_n(s)} \log \frac{c_n(x,s)}{c_n(s)}$, this completes the proof.