1 Hint

This document provides a hint regarding how to solve the first question of Homework 5.

The first step of applying variable elimination in a Bayesian Network is to find a representation of the network factors that support the basic operations of marginalization and multiplication. In the case of discrete variables, we saw that simple tabular factors were sufficient for this purpose. In the case of continuous variables, things are not so simple. In the special case where all variables have normal distributions or linear Gaussian distributions, the so-called Canonical form is sufficient to support inference using variable elimination.

Converting a single-dimensional or multi-dimensional Gaussian distribution to Canonical form is straightforward and follows directly from the definitions in the textbook (chapter 14).

However, converting a linear Gaussian CPD to Canonical form requires careful algebraic manipulation. We present here the necessary equations to convert a linear Gaussian CPD to Canonical form. You may use these equations without proof in your solution.

Let $P(X_{k+1}|X_1, ..., X_k)$ be linear Gaussian with $k{+}1$ weights $\beta_0, \beta_1, \beta_2, ..., \beta_k$. Let $\boldsymbol{\beta} = [\beta_1, ..., \beta_k]^T$ (column vector). Let σ^2 be the variance associated with X_{k+1}. This variance does not depend on $X_1, ..., X_k$.

Then $X_{k+1} = \mathcal{N}(\beta_0 + \boldsymbol{\beta}^T[X_1, ..., X_k]^T; \sigma^2)$, and the canonical form becomes a factor $C(X_{k+1}, X_1, ..., X_k; K, \mathbf{h}, g)$, where:

$$g = -\frac{1}{2}\frac{1}{\sigma^2}\beta_0^2 - \frac{1}{2}\ln(2\pi\sigma^2) \tag{1}$$

$$\mathbf{h} = \begin{bmatrix} \frac{1}{\sigma^2}\beta_0 \\ -\boldsymbol{\beta}\frac{1}{\sigma^2}\beta_0 \end{bmatrix} \tag{2}$$

$$\mathbf{K} = \begin{bmatrix} \frac{1}{\sigma^2} & -\frac{1}{\sigma^2}\boldsymbol{\beta}^T \\ -\boldsymbol{\beta}\frac{1}{\sigma^2} & \boldsymbol{\beta}\frac{1}{\sigma^2}\boldsymbol{\beta}^T \end{bmatrix} \tag{3}$$

Notice that $\boldsymbol{\beta}$ is a k-element vector, so $\mathbf{h}$ is a $k + 1 \times 1$ dimensional column vector, and K is a $k + 1 \times k + 1$ dimensional matrix. Thus, the generated factor has a scope of $k + 1$ variables.