Universidad de Cantabria

REDES BAYESIANAS NORMALES (RBN)

En una red bayesiana normal, las variables X son normales $N(\mu, \Sigma)$ con función de densidad:

$$(2\pi)^{-n/2}|\Sigma|^{-1/2} \exp\left\{-1/2(x-\mu)^T\Sigma^{-1}(x-\mu)\right\},$$

donde μ es el vector n-dimensional de medias, Σ es la matriz $n\times n$ de covarianzas, $|\Sigma|$ es el determinante de Σ, y μ^T es la traspuesta de μ.
Como en toda red Bayesiana, la función de densidad conjunta puede escribirse como producto de las funciones de probabilidad condicionadas:

$$f(x_i|\pi_i) \sim N\left(\mu_i + \sum_{j=1}^{i-1} \beta_{ij}(x_j - \mu_j),\ v_i\right), \qquad (1)$$

donde β_{ij} es el coeficiente de regresión de X_j en la regresión de X_i sobre los padres, Π_i, de X_i y

$$v_i = \Sigma_i - \Sigma_{i\Pi_i}\Sigma_{\Pi_i}^{-1}\Sigma_{i\Pi_i}^T$$

es la varianza condicionada de X_i, dado $\Pi_i = \pi_i$, donde Σ_i es la varianza incondicional de X_i, $\Sigma_{i\Pi_i}$ son las covarianzas entre X_i y las variables de Π_i, y Σ_{Π_i} es la matriz de covarianzas de Π_i.

REDES BAYESIANAS NORMALES

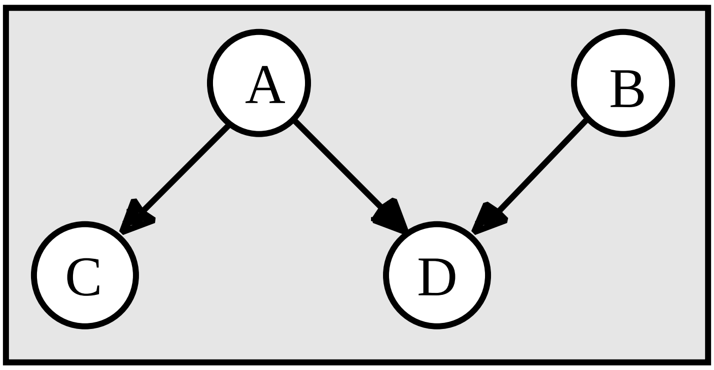
Obsérvese que β_{ij} mide el grado de relación existente entre las variables X_i y X_j. Si $\beta_{ij}=0$, entonces X_j no será un padre de X_i.
Una red Bayesiana normal depende de los parámetros $\{\mu_1,\ldots,\mu_n\}$, $\{v_1,\ldots,v_n\}$, y $\{\beta_{ij}\,|\,j<i\}$.
Una función de probabilidad normal puede definirse de forma alternativa mediante su vector de medias μ y su matriz de precisión $W=\Sigma^{-1}$. La transformación recursiva para pasar de $\{v_1,\ldots,v_n\}$ y $\{\beta_{ij}:j<i\}$ a W es:

$$W(i+1)=\begin{pmatrix} W(i)+\dfrac{\beta_{i+1}\beta_{i+1}^T}{v_{i+1}} & \dfrac{-\beta_{i+1}}{v_{i+1}} \\ \dfrac{-\beta_{i+1}^T}{v_{i+1}} & \dfrac{1}{v_{i+1}} \end{pmatrix}, \qquad (2)$$

donde $W(1)=1/v_1$, $W(i)$ es la matriz superior izquierda $i\times i$ de W y β_i es el vector columna $\{\beta_{ij}:j<i\}$
Por tanto, se tienen dos representaciones alternativas de la función de probabilidad de una red Bayesiana normal.

Universidad de Cantabria

REDES BAYESIANAS NORMALES



Considérese la red bayesiana de la figura sobre $\{A, B, C, D\}$ con $f(a, b, c, d) \sim N(\mu, \Sigma)$, que puede factorizarse como:

$$f(a, b, c, d) = f(a)f(b)f(c|a)f(d|a, b), \qquad (3)$$

donde

$$f(a) \sim N(\mu_A, v_A),$$
$$f(b) \sim N(\mu_B, v_B),$$
$$f(c|a) \sim N(\mu_C + \beta_{CA}(a - \mu_A), v_C),$$
$$f(d|a, b) \sim N(\mu_D + \beta_{DA}(a - \mu_A) + \beta_{DB}(b - \mu_B), v_D).$$

Los parámetros de esta representación son

$$\{\mu_A, \mu_B, \mu_C, \mu_D\},$$
$$\{v_A, v_B, v_C, v_D\}, y$$
$$\{\beta_{CA}, \beta_{DA}, \beta_{DB}\}.$$

Una representación alternativa la da W:

$$\begin{pmatrix} \frac{1}{v_A}+\frac{\beta_{CA}^2}{v_C}+\frac{\beta_{DA}^2}{v_D} & \frac{\beta_{DA}\beta_{DB}}{v_D} & -\frac{\beta_{CA}}{v_C} & -\frac{\beta_{DA}}{v_D} \\ \frac{\beta_{DA}\beta_{DB}}{v_D} & \frac{1}{v_B}+\frac{\beta_{DB}^2}{v_D} & 0 & -\frac{\beta_{DB}}{v_D} \\ -\frac{\beta_{CA}}{v_C} & 0 & \frac{1}{v_C} & 0 \\ -\frac{\beta_{DA}}{v_D} & -\frac{\beta_{DB}}{v_D} & 0 & \frac{1}{v_D} \end{pmatrix}.$$

La matriz de covarianzas resulta:

$$\begin{pmatrix} v_A & 0 & \beta_{CA}v_A & \beta_{DA}v_A \\ 0 & v_B & 0 & \beta_{DB}v_B \\ \beta_{CA}v_A & 0 & \beta_{CA}^2 v_A + v_C & \beta_{CA}\beta_{DA}v_A \\ \beta_{DA}v_A & \beta_{DB}v_B & \beta_{CA}\beta_{DA}v_A & \beta_{DA}^2 v_A + \beta_{DB}^2 v_B + v_D \end{pmatrix}.$$

Considerando nulas las medias, uno las varianzas, $\beta_{CA}=1$, $\beta_{DA}=0.2$ y $\beta_{DB}=0.8$, se tiene

$$\Sigma = \begin{pmatrix} 1.0 & 0.0 & 1.0 & 0.20 \\ 0.0 & 1.0 & 0.0 & 0.80 \\ 1.0 & 0.0 & 2.0 & 0.20 \\ 0.2 & 0.8 & 0.2 & 1.68 \end{pmatrix}. \tag{4}$$

Teorema 1 Distribución normal multivariada. *Sean Y y Z dos conjuntos de variables aleatorias con función de distribución normal multivariada cuyo vector de medias y matriz de covarianzas son*

$$\mu = \begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix} \quad \text{y} \quad \Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix},$$

donde μ_Y y Σ_{YY} son el vector de medias y la matriz de covarianzas de Y, μ_Z y Σ_{ZZ} son el vector de medias y la matriz de covarianzas de Z, y Σ_{YZ} es la matriz de covarianzas de Y y Z. Entonces, la función de probabilidad condicionada de Y dado $Z = z$ es una función normal multivariada con vector de medias $\mu_{Y|Z=z}$ y matriz de covarianzas $\Sigma_{Y|Z=z}$, donde

$$\mu_{Y|Z=z} = \mu_Y + \Sigma_{YZ}\Sigma_{ZZ}^{-1}(z - \mu_Z), \quad (5)$$

$$\Sigma_{Y|Z=z} = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY}. \quad (6)$$

Es conveniente actualizar de uno en uno los nodos evidenciales.

```
(* Definición de la función de probabilidad *)
M={0,0,0,0};
V={{1.0, 0.0, 1.0, 0.2},
   {0.0, 1.0, 0.0, 0.8},
   {1.0, 0.0, 2.0, 0.2},
   {0.2, 0.8, 0.2, 1.68}};
(* Nodos y evidencia *)
X={A,B,C,D};Ev={A,B,C};ev={1,3,2};
(* Actualización de M y V *)
NewM=Transpose[List[M]];
NewV=V;
For[k=1, k<=Length[Ev], k++,
(* Posición del elemento i-ésimo de E[[k]] en X *)
    i=Position[X,Ev[[k]]][[1,1]];
    My=Delete[NewM,i];
    Mz=NewM[[i,1]];
    Vy=Transpose[Delete[Transpose[Delete[NewV,i]],i]];
    Vz=NewV[[i,i]];
    Vyz=Transpose[List[Delete[NewV[[i]],i]]];
    NewM=My+(1/Vz)*(ev[[k]]-Mz)*Vyz;
    NewV=Vy-(1/Vz)*Vyz.Transpose[Vyz];
(* Eliminar el elemento i-ésimo *)
    X=Delete[X,i];
(* Imprimir los resultados *)
    Print["Iteracion = ",k];Print["Nodos restantes = ",X];
    Print["M = ",Together[NewM]];
    Print["V = ",Together[NewV]];
    Print["--------------------"];]
```

Teorema 2 *Considérese una red Bayesiana normal definida en* $X = \{X_1, \ldots, X_n\}$ *con vector de media* μ *y matriz de covarianzas* Σ*. Particionemos* X*,* μ *y* Σ *en la forma* $X = \{Y, Z\}$*,*

$$\mu = \begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix} \quad \text{y} \quad \Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix},$$

donde μ_Y *y* Σ_{YY} *son el vector de medias y la matriz de covarianzas de* Y*,y* Σ_{YZ} *es la matriz de covarianzas de* Y *y* Z*. Supóngase que* Z *es el conjunto de nodos evidenciales. Entonces la distribución condicional de cualquier variable* $X_i \in Y$ *dada* Z*, es normal, cuya media y varianza son cocientes de funciones polinomiales en las variables evidenciales y los correspondientes parámetros en* μ *y* Σ*. Los polinomios son como mucho de grado uno en las variables condicionantes, en los parámetros de medias y varianzas, y son de grado dos en los parámetros de covarianzas. El polinomio del denominador es el mismo para todos los nodos.*

```
(* Vector de medias y matriz de covarianzas *)
mean={m,0,0,1,0};
var={{a,1,1,2,1},{1,2,1,-1,1},{1,1,b,-1,c},
    {2,-1,-1,4,-2},{1,1,c,-2,6}};
(* Orden de conocimiento de la evidencia *)
evidencia={3,5}
(* Probabilidades Marginales *)
For[k=0,k<=Length[evidencia],k++,
 For[i=1,i<=Length[mean],i++,
  If[MemberQ[Take[evidencia,k],i],
   cmean=x[i];
   cvar=0,
   meany=mean[[i]];
   meanz=Tabla[{mean[[evidencia[[j]]]]},{j,1,k}];
   vary=var[[i,i]];
   If[k==0,
    cmean=Together[meany];
    cvar=Together[vary],
    varz=Tabla[Tabla[
        var[[evidencia[[t]],evidencia[[j]]]],
      {t,1,k}],{j,1,k}];
    covaryz=Tabla[{var[[evidencia[[t]]]][[i]]},{t,1,k}];
    zaux=Tabla[{x[evidencia[[t]]]},{t,1,k}];
    aux=Inverse[varz];
    cmean=meany+Transpose[covaryz].aux.(zaux-meanz);
    cvar=vary-Transpose[covaryz].aux.covaryz;]];
  Print["Nodos evidenciales ",k," Nodo ",i];
  Print["Media ",Together[cmean],"Var ",Together[cvar]];]]
```

Universidad de Cantabria

EJEMPLO DE PROPAGACION SIMBOLICA

Considérese el conjunto de variables

$$X = \{X_1, \ldots, X_5\}$$

con vector de medias y matriz de covarianzas

$$\mu = \begin{pmatrix} m \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \Sigma = \begin{pmatrix} a & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & -1 & 1 \\ 1 & 1 & b & -1 & c \\ 2 & -1 & -1 & 4 & -2 \\ 1 & 1 & c & -2 & 6 \end{pmatrix}. \qquad (7)$$

Nótese que la media de X_1, la varianza de X_1 y X_3, y la covarianza de X_3 y X_5 están especificadas en forma simbólica.
Se desea conocer las probabilidades marginales iniciales de los nodos (no evidencia) y las probabilidades condicionales de los nodos dadas cada una de las evidencias $\{X_3 = x_3\}$ y $\{X_3 = x_3, X_5 = x_5\}$. Las medias y varianzas condicionadas son expresiones racionales, es decir, cocientes de polinomios en los parámetros.

Universidad de Cantabria

EJEMPLO DE PROPAGACION SIMBOLICA

No Evidencia		
X_i	Media	Varianza
X_1	m	a
X_2	0	2
X_3	0	b
X_4	1	4
X_5	0	6
Evidencia $X_3 = x_3$		
X_1	$(bm + x_3)/b$	$(ab - 1)/b$
X_2	$(x_3)/b$	$(2b - 1)/b$
X_3	x_3	0
X_4	$(b - x_3)/b$	$(4b - 1)/b$
X_5	$(cx_3)/b$	$(6b - c^2)/b$
Evidencia $X_3 = x_3$ y $X_5 = x_5$		
X_1	$\dfrac{6bm - c^2m + (6-c)x_3 + (b-c)x_5}{6b - c^2}$	$\dfrac{6ab + 2c - ac^2 - b - 6}{6b - c^2}$
X_2	$\dfrac{(6-c)x_3 + (b-c)x_5}{6b - c^2}$	$\dfrac{11b + 2c - 2c^2 - 6}{6b - c^2}$
X_3	x_3	0
X_4	$\dfrac{6b - c^2 + (2c-6)x_3 + (c-2b)x_5}{6b - c^2}$	$\dfrac{20b + 4c - 4c^2 - 6}{6b - c^2}$
X_5	x_5	0