7 | Factores

Luis Alfredo Lizárraga Santos
Verónica Esther Arriola Ríos

Objetivo

Que el alumno implemente una clase `Factor` para familiarizarse con las operaciones entre factores que se utilizan para realizar inferencia en redes Bayesianas.

Introducción

Un factor es una estructura matemática que nos ayudará a sistematizar las diversas operaciones requeridas para operar con distribuciones de probabilidad discretas. Por ello, en esta práctica, el alumno se familiarizará con estas estructuras e implementará sus operaciones en un lenguaje de programación. Posteriormente, la biblioteca creada será la herramienta con la cual se resolverán problemas de aplicación diversos.

Definición 7.1: Factor

"Sea D un conjunto de variables aleatorias [y $Val(D)$ una asignación de valores a estas variables]. Se define un Factor ϕ como una función de $Val(D)$ a $\mathbb{R}$. Un factor es no-negativo si todas sus entradas son no negativas. Al conjunto de variables D se le llama *alcance del factor* y se denota como $Alcance(\phi)$"

(Koller y Friedman 2009, pág. 104).

Tomando como base la definición de Koller y Friedman, podemos realizar un parafraseo y afirmar que un factor es una estructura matemática definida sobre un conjunto de variables donde cada variable puede tomar valores de su dominio, el cual debe ser finito. El factor asocia un número real a cada posible asignación de valores de esas variables.

A	$\phi(A)$
0	.3
1	.7

(a) Factor A

B	$\phi(B)$
0	.6
1	.4

(b) Factor B

C	$\phi(C)$
0	.2
1	.8

(c) Factor C

Figura 7.1 Factores sobre variables aleatorias binarias.

En cuanto a las operaciones de factores que se explicarán en esta práctica, nos basamos en las operaciones que presentan Koller y Friedman: multiplicación, reducción, marginalización y normalización.

Utilizaremos a los factores para realizar operaciones con distribuciones de probabilidad, sin embargo es importante notar que, debido a la misma definición de factor, el resultado de cada operación entre factores no siempre será una medida de probabilidad, pues esta limita su valores al intervalo $[0, 1]$, mientras que los factores ocupan todo $\mathbb{R}$. Adicionalmente, sus operaciones no siempre incluyen las normalizaciones requeridas por las distribuciones de probabilidad; pero esto se convertirá en una ventaja para nosotros, como veremos más adelante.

Operaciones entre Factores

Multiplicación

Definición 7.2: Multiplicación

"Sean X, Y y Z tres conjuntos disjuntos de variables, y sean $\phi_1(X, Y)$ y $\phi_2(Y, Z)$ dos factores. Se define el factor producto $\phi_1 \times \phi_2$ como un factor $\psi : Val(X, Y, Z) \mapsto \mathbb{R}$ de la forma siguiente:

$$\psi(X, Y, Z) = \phi_1(X, Y) \cdot \phi_2(Y, Z)$$

"

(Koller y Friedman 2009, pág. 107).

La operación de multiplicación es sencilla. Si los conjuntos de variables de los factores a multiplicar no tienen elementos en común, se multiplica cada entrada del factor A por cada entrada del Factor B. El alcance del factor resultante es la unión de los alcances de los factores a multiplicar (Koller y Friedman 2009, pág. 107).

Por ejemplo: Tenemos los tres factores A, B y C de la Figura 7.1. Para obtener el factor $AB = A \times B$, se multiplicaría el renglón $A = 0$ con $B = 0$, luego $A = 0$ con $B = 1$, $A = 1$ con $B = 0$ y por último $A = 1$ con $B = 1$ como en la Figura 7.2.

A	B	ϕ(A,B)
0	0	(.3)*(.6) = 0.18
0	1	(.3)*(.4) = 0.12
1	0	(.7)*(.6) = 0.42
1	1	(.7)*(.4) = 0.28

Figura 7.2 Factor AB

Si el alcance de ambos factores tiene variables en común, se debe asegurar que tengan el mismo valor en cada renglón por multiplicar en ambos factores. Por ejemplo, si se tienen los factores AB y AC, al momento de multiplicar el renglón $\{A = 0, B = 0\}$ se debe seleccionar los renglones donde $A = 0$ en el factor AC, estos son $\{A = 0, C = 0\}$ y $\{A = 0, C = 1\}$, como muestran Figura 7.3 y Figura 7.4.

A	B	ϕ(A,B)
0	0	(.3)*(.6) = 0.18
0	1	(.3)*(.4) = 0.12
1	0	(.7)*(.6) = 0.42
1	1	(.7)*(.4) = 0.28

(a) Factor $AB = A \times B$

A	C	ϕ(A,C)
0	0	(.3)*(.2) = 0.6
0	1	(.3)*(.8) = 0.24
1	0	(.7)*(.2) = 0.14
1	1	(.7)*(.8) = 0.56

(b) Factor $AC = A \times C$

Figura 7.3 Multiplicandos.

A	B	C	ϕ(A,B,C)
0	0	0	[(.3)*(.6)]*[(.3)*(.2)] = 0.0108
0	0	1	[(.3)*(.6)]*[(.3)*(.8)] = 0.0432
0	1	0	[(.3)*(.4)]*[(.3)*(.2)] = 0.0072
0	1	1	[(.3)*(.4)]*[(.3)*(.8)] = 0.0288
1	0	0	[(.7)*(.6)]*[(.7)*(.2)] = 0.0588
1	0	1	[(.7)*(.6)]*[(.7)*(.8)] = 0.2352
1	1	0	[(.7)*(.4)]*[(.7)*(.2)] = 0.0392
1	1	1	[(.7)*(.4)]*[(.7)*(.8)] = 0.1568

Figura 7.4 Factor $AB \times AC$

Es importante hacer notar que al multiplicar los factores AB y AC no se estaría obteniendo la probabilidad conjunta de A, B y C, si no alguna otra función $\phi(A, B, C)$.

Reducción

Definición 7.3: Reducción

"Sea $\phi(Y)$ un factor y $U = u$ una asignación para $U \subseteq Y$. Se define la reducción de un factor ϕ al contexto $U = u$, denotado como $\phi[U = u]$ (y abreviado como $\phi[u]$), como un factor con alcance $Y' = Y - U$ de tal forma que

$$\phi[u](y') = \phi(y', u)$$

"

(Koller y Friedman 2009, pág. 111).

La operación de reducción consiste en seleccionar un valor de alguna variable del factor y sólo tomar los renglones que cumplen con el valor dado de la variable. Por ejemplo: Se tiene el factor AB de la Figura 7.5.

A	B	ϕ(A,B)
0	0	.18
0	1	.12
1	0	.42
1	1	.28

Figura 7.5 Factor AB

Si se desea reducir con $A = 0$, el resultado es el factor en la Figura 7.6.

B	$\phi(B)$
0	.18
1	.12

Figura 7.6 Factor B

Normalización

Para normalizar un factor, esto es, que los valores que toma el factor sumen 1, basta con sumar todos los valores asociados a las asignaciones y dividir cada uno entre esta suma. Por ejemplo: Tenemos el factor $\phi(B)$ obtenido en la Figura 7.6, la suma de sus posibles valores es .3, entonces el factor normalizado queda como en la Figura 7.7.

A	B	$\phi(B)$
0	0	(.18/.3)= .6
0	1	(.12/.3)= .4

Figura 7.7 Factor B normalizado

Marginalización

"Sea X un conjunto de variables y sea $Y \notin X$ una variable y $\phi(X,Y)$ un factor. Se define la marginalización de Y en ϕ, denotado como $\sum_Y \phi$, como un factor ψ sobre X de tal forma que

$$\psi(X) = \sum_Y \phi(X,Y)$$

" (Koller y Friedman 2009, pág. 297).

La operación de marginalización consiste en tomar la variable a marginalizar, sumar los valores en los renglones en que cambia su valor pero el de las demás variables no, y asignar esta suma al renglón correspondiente de las variables restantes. Por ejemplo: tenemos el factor AB y deseamos marginalizar la variable B. Entonces, tomamos los renglones donde A=0 y los sumamos [Figura 7.8], tomamos los renglones donde A=1 y los sumamos [Figura 7.9].

	A	B	ϕ(A,B)
$\rightarrow$	0	0	.18
$\rightarrow$	0	1	.12
	1	0	.42
	1	1	.28

(a) Factor AB

	A	ϕ(A)
$\rightarrow$	0	(.18)+(.12)

(b) Factor A

Figura 7.8 Paso 1

	A	B	ϕ(A,B)
	0	0	.18
	0	1	.12
$\rightarrow$	1	0	.42
$\rightarrow$	1	1	.28

(a) Factor AB

	A	ϕ(A)
	0	(.18)+(.12)
$\rightarrow$	1	(.42)+(.28)

(b) Factor A

Figura 7.9 Paso 2

Desarrollo e implementación

La práctica consiste en crear una clase `Factor` que implemente las operaciones de multiplicación, reducción y normalización de factores y marginalización de variables. Todo esto utilizando el lenguaje de programación `Python`. Para guiar tu implementación, el código auxiliar de esta práctica consiste en un `script` de uso llamado `PruebaFactores.py`, que deberá funcionar correctamente utilizando tu clase, la cual deberá estar implementada en un archivo `Factores.py` al lado de éste.

Implementación

1. Crear una clase `Variable`, con los siguientes atributos: `nombre` y `valores_posibles` y sobrescribir el método `__str__`.

2. Crear una clase `Factor` que contenga los atributos:

   - `alcance`: una lista de objetos de clase `Variable`.
   - `valores`: una lista de valores asociados a cada renglón.

3. Programar un método auxiliar `_generar_tabla_de_valores`, que sólo ejecutará su algoritmo la primera vez que sea invocado y su uso primordialmente será para poder imprimir en pantalla el contenido del factor, es decir, servirá para sobreescrbir el método `__str__` de la clase `Factor`.

   Este método deberá generar, de forma explícita, la tabla de asignaciones posibles a las variables en el alcance del factor. Estas combinaciones serán listadas en formato semejante a un reloj digital, de modo que la última variable en el atributo `alcance` sea la que varíe su valor más rápido. Los valores posibles se irán listando en el orden en que aparecen en el atributo `valores_posibles` del objeto de tipo `Variable`.

   Observa que esta implementación no exige que los valores de la variables sean números, como indicaría la definición formal, pues en cualquier modo esto no es requisito para el correcto funcionamiento de las operaciones. Esta generalización podrá ser utilizada para visualizar resultados más rápidos de interpretar.

   Por ejemplo, para las variables:

   - `Letras = ['a', 'b']`
   - `Útiles = ['cuaderno', 'lápiz', 'goma']`
   - `Números = ['2', '8']`

   La tabla de valores posibles se extiende como en la Figura 7.10.

   Usando esto, sobreescrbir el método `__str__`.

Índice	Letras	Útiles	Números
0	a	cuaderno	2
1	a	cuaderno	8
2	a	lápiz	2
3	a	lápiz	8
4	a	goma	2
5	a	goma	8
6	b	cuaderno	2
7	b	cuaderno	8
8	b	lápiz	2
9	b	lápiz	8
10	b	goma	2
11	b	goma	8

Figura 7.10 Tabla de valores

TIP: Como tu método debe funcionar para cualquier número de variables en el alcance generar esta tabla requerirá una forma creativa de anidar ciclos `for`. En algún lugar necesitarás iterar sobre la misma tabla de valores que estás generando.

4. Programar un método auxiliar que reciba como parámetro un diccionario, cuyas llaves sean variables y los valores sean el valor asignado a cada una. Este método deberá obtener el índice en la tabla de valores que corresponda a esta asignación. Observa que para obtener el índice no es necesario haber desarrollado la tabla explícitamente, se utiliza un polinomio de direccionamiento.

   Por ejemplo, el polinomio de direccionamiento para un factor de 3 variables (A, B y C) se vería así:

   $$índice = (pos(a) * |B| * |C|) + (pos(b) * |C|) + pos(c)$$

   donde $pos(a)$ es la posición del valor asociado a la variable A en su lista de valores posibles y $|A|$ es el tamaño de esta lista.

   En su forma más general el polinomio de direccionamiento tiene la forma:

   $$índice = \sum_{Var \in D} \left(pos(Val(Var)) \prod_{Var \in > D} |Var| \right)$$

   donde $Var \in > D$ son todas las variables a la derecha de Var en la lista de `valores_posibles`.

   TIP: Pueden factorizar los términos comunes (los tamaños de las listas) e implementar una función auxiliar que calcule recursivamente el valor del polinomio de direccionamiento.

5. Implementa las operaciones: multiplicación, reducción, normalización y marginalización.

- En el caso de la multiplicación, considera también el caso de multiplicar por un escalar. Si el parámetro recibido es un escalar, el factor resultado tendrá el mismo alcance y sus valores serán multiplicados todos por el escalar.

- Para la marginalización, si se pide marginalizar la única variable del factor, suma los renglones y devuelve el escalar.

- En la reducción, si se reduce la única variable, devuelve el valor del factor que corresponda al valor indicado de la variable.

TIP: Crea primero el factor resultado con lista de variables en su alcance. Para cada renglón en la tabla de valores de este factor, encuentra los renglones relevantes en los operandos utilizando el método auxiliar mencionado en el punto anterior y realiza la operación correspondiente.

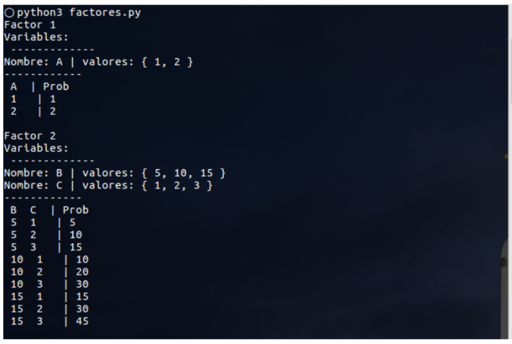
Esta práctica deberá ser implementada usando `Python`.

Requisitos y resultados

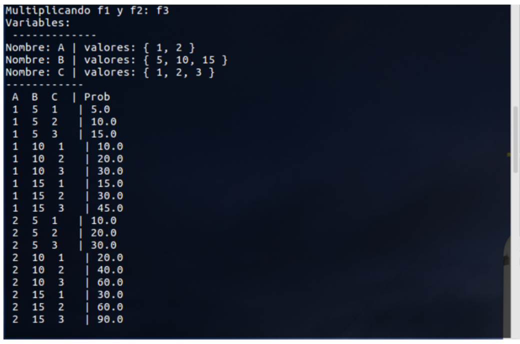
Deberán hacer casos de prueba para marginalización, reducción, normalización y multiplicación de factores. Basta con incluir un pequeño *script* con sus casos de prueba. No es necesario que lo hagan a prueba de todo, pero sí que funcionen cuando se cumplen los prerrequisitos para cada operación.

Para evaluar y calificar la práctica es necesario que se implementen todos los métodos mencionados e indicados, respetando las especificaciones de estilo y documentación del lenguaje de programación que usarán. Es completamente válido utilizar bibliotecas adicionales si lo consideran necesario, así como la creación y uso de sus propios métodos auxiliares si lo desean.

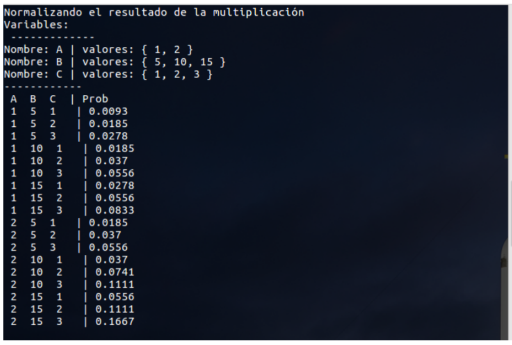
En la figura 7.11 pueden apreciar el resultado de crear 3 factores y ejecutar las operaciones de multiplicación, normalización, reducción y marginalización entre ellos.



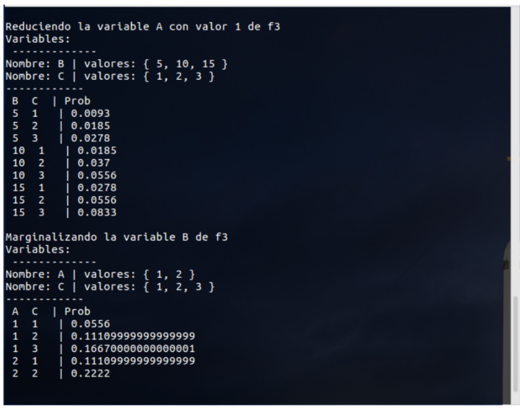
(a) Creación de factores f1 y f2



(b) Multiplicando f1 y f2, para obtener f3



(c) Normalizando f3



(d) Reduciendo y marginalizando

Figura 7.11 Ejecución del código solución: crea dos factores, los multiplica, normaliza, reduce y marginaliza

No olviden documentar y comentar su código.