Facultad de Ciencias

ESPACIOS LINEALES TROPICALES
(TROPICAL LINEAR SPACES)

Trabajo de Fin de Grado
para acceder al

GRADO EN MATEMÁTICAS

Autor: Jorge Seco Lavid

Director: Luis Felipe Tabera Alonso

Febrero - 2017

Resumen

El objetivo de este trabajo es definir los espacios lineales tropicales. Para ello, primero analizaremos el caso clásico. Veremos la Grassmanniana, una variedad en la que cada punto se corresponde con un subespacio lineal. Realizaremos una introducción general de la geometría tropical y estudiaremos el análogo tropical de la Grassmanniana, llamado Dressiano. Con estas herramientas, estamos preparados para definir espacio lineal tropical. Como aplicación, veremos los árboles filogenéticos, de los cuales daremos las definiciones básicas y mostraremos un caso con datos reales de ADN mitocondrial humano.

Palabras clave: Grassmanniana, Geometría Tropical, Dressiano, Árboles filogenéticos

Abstract

In this report we define tropical linear spaces. For this, we start analyzing the classical case, defining a variety called Grassmannian, where each point corresponds to a linear subspace. We introduce tropical geometry and we study his tropical analogue, called the Dressian. With these tools, we are ready to define tropical linear spaces. As an aplication, phylogenetic trees are defined and we show an example with real data of human mitochondrial DNA.

Key words: Grassmannian, Tropical Geometry, Dressian, Phylogenetic tree.

Índice general

Capítulo 1

Introducción

En el este Trabajo de Fin de Grado estudiaremos la geometría tropical. Se trata de un área reciente de las matemáticas que ha sufrido un rápido desarrollo desde principios del siglo XXI. El adjetivo tropical se trata de un “homenaje”del informático francés Jean-Eric Pin al brasileño Imre Simon, autor de uno de los primeros artículos que llevan la palabra tropical “On semigroups of matrices over the tropical semiring”. Es decir, que no hay ningún significado profundo detrás de este curioso adjetivo, sino que simplemente indica la visión de Jean-Eric sobre Brasil. Se han hecho conexiones profundas entre la geometría tropical y muchas ramas de la matemática pura y aplicada. La geometría tropical se ha convertido en un campo grande, y sólo abordaremos una pequeña introducción. Se desarrolla sobre el semianillo tropical $\mathbb{T} = \{\mathbb{R} \cup \infty, \oplus, \odot\}$ con las operaciones habituales de la adición y multiplicación sustituidas por el mínimo y la suma respectivamente.

Trataremos de estudiar el análogo tropical del espacio lineal clásico, pero nos encontraremos con que en algunas situaciones los objetos no se comportan como nos gustaría. Para solucionarlo introduciremos dos herramientas: la Grassmanniana y el Matroide.

La Grassmanniana $Gr(n, m)$ es un espacio que parametriza todos los subespacios lineales de dimensión n de un espacio lineal de dimension m, es decir, que cada punto del Grassmanniano corresponde a una variedad algebraica. Una matroide es una estructura que generaliza el concepto de independencia lineal en espacios vectoriales.

Estudiaremos estos dos conceptos antes de entrar en el álgebra tropical. Veremos sus propiedades básicas, su aritmética y los polinomios tropicales. Al intentar definir el espacio vectorial tropical nos encontraremos con que la definición natural “no es buena”. Para hallar una definición buena tropicalizaremos la Grassmanniana clásica que tiene le estudiaremos porque tiene mejores propiedades combinatorias.

Por último, introduciremos los árboles filogenéticos, que son árboles de m hojas y sin vértices de grado 2. Estos surgen sobretodo en biología (aunque no es el único campo donde se han usado), donde las etiquetas representan diferentes especies y la estructura de árbol registra su historia evolutiva. Relacionaremos la Grassmanniana Tropical con el conjunto de los árboles filogenéticos.

Capítulo 2

Grassmaniana clásica

Los espacios de moduli son objetos fundamentales en la geometría algebraica. Estos espacios parametrizan familias de variedades. Cada punto de un espacio de moduli corresponde a una variedad algebraica diferente en la familia de interés. Un caso básico es la familia de subespacios n-dimensionales del espacio vectorial $\mathbb{K}^m$. Esta familia es parametrizada por la grassmaniana $Gr(n, m)$, donde cada punto representa un n-plano de $\mathbb{K}^m$. Se trata una variedad proyectiva de dimensión $n(m-n)$.

En las próximas páginas, vamos a dar una serie de proposiciones y definiciones para construir la Grassmanniana.

Definición 2.1. *La Grassmanniana $Gr(n, m)$ se define como el conjunto de n-planos de $\mathbb{K}^m$.*

Proposición 2.2. *Todo n-plano de $\mathbb{K}^m$ se puede representar con una matriz $n \times m$ de rango n.*

Demostración. Sea H un n-plano de $\mathbb{K}^m$. Sea $\beta = \{\beta_1, ..., \beta_n\}$ una base cualquiera de H. Tomemos los vectores de β como filas de M. Como son n vectores de m coordenadas y además son independientes, tenemos una matriz $n \times m$ de rango n.

□

Proposición 2.3. *Toda matriz $n \times m$ ($n \leq m$) de rango n define un n-plano de $\mathbb{K}^m$.*

Demostración. Sea A una matriz $n \times m$ de rango n. Tomemos las n filas de A como vectores β_i de m coordenadas. Como el rango de A es n, los vectores β_i son independientes y forman una base de un n-plano de $\mathbb{K}^m$. □

Definición 2.4. *Llamamos $\mathcal{A}(n,m)$ al subconjunto de las matrices $n \times m$, de rango n, de la forma:*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

Definición 2.5. *Sea $A \in \mathcal{A}(n,m)$. El espacio generado por sus filas se llama espacios de filas de A y se escribe $rs(A)$ (rowspan(A)).*

Proposición 2.6. *Dos matrices A y $B \in \mathcal{A}(n,m)$ definen el mismo n-plano si y sólo si $\exists C \in GL_n(\mathbb{K})$ con $CA = B$.*

Demostración. "$\Longleftarrow$"
Sean A y B matrices de $\mathcal{A}(n,m)$, y $C \in GL_n(\mathbb{K})$:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}; \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}; \quad C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}$$

Supongamos que: $CA = B$, y sea B_i la i-ésima fila de B:

$B_i = (C_i)A \Rightarrow B_i \in rs(A)$, $\forall i \in \{1, ..., m\}$. Por tanto $rs(B) \subseteq rs(A)$. La igualdad se obtiene porque $dim(rs(B)) = dim(rs(A)) = n$.

"$\Longrightarrow$"
Sean A y $B \in \mathcal{A}(n,m)$ y supongamos que definen el mismo n-plano. Sea B_i una fila cualquiera de B. Como A y B definen el mismo n-plano, podemos escribir $B_i = C_i A$.
Esto se tiene para todo $i \in \{1, ..., n\}$. Así, construimos una matriz C de tamaño $n \times n$, que se trata de la matriz de cambio de base de B a A y, por lo tanto su determinante es distinto de cero, por lo que $C \in GL_n(\mathbb{K})$

□

Por la proposición 2.3 todo elemento de la Grassmanniana $Gr(n,m)$ se puede representar por medio de una matriz $A \in \mathcal{A}(n,m)$.

$$A = \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}$$

Dos matrices A y $B \in \mathcal{A}(n,m)$ representan el mismo elemento de $Gr(n,m)$ si y sólo si $B = CA$, para algún $C \in GL_n(\mathbb{K})$.

Ejemplo 2.7. *Sea $A \in \mathbb{K}^6$ la siguiente matriz:*

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Sea $C \in GL_3(\mathbb{K})$:

$$C = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$

Por lo tanto:

$$CA = \begin{pmatrix} 4 & 8 & 2 & 2 & 0 & 6 \\ 2 & 4 & 3 & 2 & -1 & 4 \\ 0 & 0 & 1 & 3 & 2 & 3 \end{pmatrix}$$

Veamos que los puntos A y CA representan el mismo elemento de $Gr(n,m)$: Para ello calculo $dim(rs(A))$, $dim(rs(CA))$ y $dim(rs(A) + rs(CA))$ y veo que son la misma:
Sea $p_{A_{135}}$ el menor de A formado por las columnas 1, 3 y 5, y sea $p_{CA_{156}}$ el menor de CA formado por las columnas 1, 5 y 6. Es fácil comprobar que sus determinantes son distintos de 0:

$$p_{A_{135}} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1 \neq 0 \quad ; \quad p_{CA_{156}} = \begin{vmatrix} 4 & 0 & 6 \\ 2 & -1 & 4 \\ 0 & 2 & 3 \end{vmatrix} = -20 \neq 0$$

De este modo queda comprobado que $dim(rs(A)) = dim(rs(CA)) = 3$
Si comprobamos que $dim(rs(A) + rs(CA)) = 3$, entonces queda demostrado que A y CA generan el mismo subespacio vectorial y que por tanto representan el mismo elemento de $Gr(n,m)$:

$$rs(A)+rs(CA)=\begin{pmatrix}0&0&1&1&0&1\\0&0&0&1&1&1\\1&2&0&0&0&1\\4&8&2&2&0&6\\2&4&3&2&-1&4\\0&0&1&3&2&3\end{pmatrix}$$

Calculamos la dimensión de $A+CA$ utilizando el algoritmo de eliminación Gaussiana, operando por sus filas, para encontrar sus dependencias:

$$\begin{pmatrix}1&2&0&0&0&1\\0&0&1&1&0&1\\0&0&0&1&1&1\\4&8&2&2&0&6\\2&4&3&2&-1&4\\0&0&1&3&2&3\end{pmatrix}\to\begin{pmatrix}1&2&0&0&0&1\\0&0&1&1&0&1\\0&0&0&1&1&1\\0&0&2&2&0&2\\0&0&3&2&-1&2\\0&0&1&3&2&3\end{pmatrix}$$

$$\to\begin{pmatrix}1&2&0&0&0&1\\0&0&1&1&0&1\\0&0&0&1&1&1\\0&0&0&0&0&0\\0&0&0&-1&-1&-1\\0&0&0&2&2&2\end{pmatrix}\to\begin{pmatrix}1&2&0&0&0&1\\0&0&1&1&0&1\\0&0&0&1&1&1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}$$

Por lo tanto $dim(rs(A)+rs(CA))=3$, así que A y CA representan el mismo elemento de $Gr(n,m)$.

Proposición 2.8. *Los n-planos de $\mathbb{K}^m$, que son los puntos de $Gr(n,m)$, están en biyección con las clases de matrices de $\mathcal{A}(n,m)$, módulo la relación de equivalencia $\sim$:*

$$A\sim B\iff \exists C\in GL_n(\mathbb{K}): A=CB$$

Demostración. Sea f la siguiente aplicación:

$$f:\mathcal{A}(n,m)\longrightarrow Gr(n,m)$$

tal que $f(A)=rs(A)$. Esta aplicación es sobreyectiva por la Proposición 2.2. Por otro lado, la Proposición 2.6 nos dice que $f(A)=f(B)$ si y sólo si $A\sim B$. Teniendo esto en cuenta, la aplicación que va del conjunto cociente a la imagen:

$$\hat{f}:\mathcal{A}(n,m)/\sim\longrightarrow Gr(n,m)$$

es biyectiva. □

Vamos a ver ahora una representación de la Grassmanniana $Gr(n, m)$ como variedad proyectiva en $\mathbb{P}^{\binom{m}{n}-1}$.

Definición 2.9. *Sea $A \in \mathcal{A}(n, m)$. Nos referimos como $A^{i_1,\ldots,i_n}$ a la matriz formada por las columnas $i_1, ..., i_n$ de A. Llamamos menor de las columnas $i_1, ..., i_n$ y lo escribimos $p_{i_1,\ldots,i_n}$ al determinante de esta matriz $A^{i_1,\ldots,i_n}$. En total hay $\binom{m}{n}$ menores. Los podemos ordenar en el órden lexicográfico de las columnas. A esta n-tupla de menores lo llamamos coordenadas de Plücker de A.*

Definición 2.10. *Sean $i_1, ..., i_m$ las columnas de una matriz $A \in \mathcal{A}(n, m)$. Sea $\phi : \mathcal{A}(n, m) \longrightarrow \mathbb{P}^{\binom{m}{n}-1}$ tal que $\phi(A) = [menor(A; i_1, ..., i_n)$ en órden lexicográfico$]$. Como $A \in \mathcal{A}(n, m)$ existe al menos un menor distinto de cero y, por tanto la aplicación está bien definida.*

Ejemplo 2.11. *Sea $A \in \mathcal{A}(n, m)$.*

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Calculemos sus coordenadas de Plücker: En la siguiente matriz la primera fila representa el índice de las columnas y la segunda fila el determinante de esos menores:

$$\begin{array}{rccccccccccccccc} & 123 & & 124 & & 125 & & 126 & & 134 & & 135 & & 136 & & 145 \\ \phi(A) = [& 0 & : & 0 & : & 0 & : & 0 & : & 1 & : & 1 & : & 1 & : & 1 \\ \\ & 146 & & 156 & & 234 & & 235 & & 236 & & 245 & & 246 & & 256 \\ & 0 & : & -12 & : & 2 & : & 2 & : & 2 & : & 0 & : & -2 & : & 0 \\ \\ & 345 & & 346 & & 356 & & 456 & & & & & & & & \\ & 0 & : & 1 & : & 1 & : & 1 &] & & & & & & & \end{array}$$

Las coordenadas distintas de cero nos indican los subconjuntos de n elementos de las columnas de la matriz A que son independientes en $\mathbb{K}^n$. En el Capítulo 3 trataremos esto de manera más amplia.

Proposición 2.12. *Sea $p \in Im(\phi)$, supongamos que $p_{i_1,\ldots,i_n} \neq 0$. Entonces existe una única matriz $D \in \mathcal{A}(n, m)$ con $\phi(D) = p$ y $D^{i_1,\ldots,i_n} = Id_n$*

Demostración. Sin pérdida de generalidad, al estar trabajando con coordenadas homogéneas, podemos suponer que $p_{i_1,...,i_n} = 1$ y que $\{i_1, ..., i_n\} = \{1, ..., n\}$ son las n primeras columnas. D debe tener la forma:

$$D = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & d_{1,n+1} & \cdots & d_{1,m} \\ 0 & 1 & \cdots & 0 & \cdots & 0 & d_{2,n+1} & \cdots & d_{2,m} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & d_{i,n+1} & \cdots & d_{i,m} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & d_{n,n+1} & \cdots & d_{n,m} \end{pmatrix}$$

Ahora, calculamos los elementos que faltan de la matriz. Por ejemplo, para la primera fila de D:

$$\pm d_{1,i} = \begin{vmatrix} 0 & 0 & \cdots & 0 & d_{1,n+1} \\ 1 & 0 & \ddots & 0 & d_{2,n+1} \\ 0 & 1 & \ddots & 0 & d_{3,n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & d_{n,n+1} \end{vmatrix} = p_{2,3,...,i}$$

Donde el signo dependerá en cada caso. Para calcular un $d_{i,j}$ cualquiera, partimos de las n primeras columnas de D (la identidad),quitamos la columna con un 1 en la fila i y añadimos la columna en la que se encuentra el elemento $d_{i,j}$:

$$\pm d_{i,j} = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{1,j} \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 & d_{2,j} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & \vdots & d_{i-1,j} \\ \vdots & \vdots & \ddots & 0 & 0 & \ddots & \vdots & d_{i,j} \\ 0 & 0 & \cdots & 0 & 1 & \cdots & \vdots & d_{i+1,j} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & d_{m,j} \end{vmatrix} = p_{1...,i-1,i+1,...j}$$

□

Ejemplo 2.13. *Veamos un ejemplo de cómo recuperar una matriz D a partir de las coordenadas de Plücker.*
Supongamos que tenemos las coordenadas de Plücker del Ejemplo 2.11:

$\phi(A) = [0:0:0:0:1:1:1:1:0:-1:2:2:2:2:0:-2:0:1:1:1]$

D será de la forma:

$$D = \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{pmatrix}$$

Refirámonos en este ejemplo al menor formado por las columnas $i_1 ..., i_n$ como $p_{i_1,...,i_n}$. Como $p_{134} = 1$, podemos suponer que las columnas 1, 3 y 4 forman la matriz identidad (como las coordenadas de Plücker son coordenadas proyectivas podríamos haber cogido tres columnas cualesquiera y multiplicar por un escalar para conseguir un 1 en la posición que deseemos). Tenemos la matriz D de la siguiente manera:

$$D = \begin{pmatrix} 1 & d_{12} & 0 & 0 & d_{15} & d_{16} \\ 0 & d_{22} & 1 & 0 & d_{25} & d_{26} \\ 0 & d_{32} & 0 & 1 & d_{35} & d_{36} \end{pmatrix}$$

Ahora, con el determinante de todos los menores y la información de algunas columnas, podemos sacar el valor de las columnas restantes:

$$p_{234} = \begin{vmatrix} d_{12} & 0 & 0 \\ d_{22} & 1 & 0 \\ d_{32} & 0 & 1 \end{vmatrix} = d_{12} = 2 \quad ; \quad p_{124} = \begin{vmatrix} 1 & d_{12} & 0 \\ 0 & d_{22} & 0 \\ 0 & d_{32} & 1 \end{vmatrix} = d_{22} = 0$$

$$p_{123} = \begin{vmatrix} 1 & d_{12} & 0 \\ 0 & d_{22} & 1 \\ 0 & d_{32} & 0 \end{vmatrix} = -d_{32} = 0 \quad ; \quad p_{345} = \begin{vmatrix} 0 & 0 & d_{15} \\ 1 & 0 & d_{25} \\ 0 & 1 & d_{25} \end{vmatrix} = d_{15} = 0$$

$$p_{145} = \begin{vmatrix} 1 & 0 & d_{15} \\ 0 & 0 & d_{25} \\ 0 & 1 & d_{35} \end{vmatrix} = -d_{25} = 1 \quad ; \quad p_{135} = \begin{vmatrix} 1 & 0 & d_{15} \\ 0 & 1 & d_{25} \\ 0 & 0 & d_{35} \end{vmatrix} = d_{35} = 1$$

$$p_{346} = \begin{vmatrix} 0 & 0 & d_{16} \\ 1 & 0 & d_{26} \\ 0 & 1 & d_{36} \end{vmatrix} = d_{16} = 1 \quad ; \quad p_{146} = \begin{vmatrix} 1 & 0 & d_{16} \\ 0 & 0 & d_{26} \\ 0 & 1 & d_{36} \end{vmatrix} = -d_{26} = 0$$

$$p_{136} = \begin{vmatrix} 1 & 0 & d_{16} \\ 0 & 1 & d_{26} \\ 0 & 0 & d_{36} \end{vmatrix} = d_{36} = 1$$

Por lo tanto, hemos conseguido la siguiente matriz:

$$D = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Puede comprobarse de manera sencilla que A y D generan el mismo subespacio vectorial. Ya que $A = CD$, con $C \in GL_3(\mathbb{K})$:

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Proposición 2.14. *Sean A, $B \in \mathcal{A}(n, m)$. Se cumple que:*

$$\phi(A) = \phi(B) \iff A \sim B$$

Demostración. "$\Longleftarrow$"
Sean $A, B \in \mathcal{A}(n, m)$ tales que $A \sim B$.
Para cada $\{i_1, ..., i_n\}$, podemos comprobar fácilmente que $B^{i_1,...,i_n} = (CA)^{i_1,...,i_n} = CA^{i_1,...,i_n}$ y, por tanto, $|B^{i_1,...,i_n}| = |C||A^{i_1,...,i_n}| \Rightarrow \phi(B) = det(C)\phi(A)$
Y por tanto, A y B definen el mismo punto proyectivo.

"$\Longrightarrow$"
Sean $A, B \in \mathcal{A}(n, m)$ dos matrices que definen el mismo punto proyectivo. Por la Proposición 2.12, sabemos que ese punto proyectivo son también las coordenadas de Plücker de una matriz $D \in \mathcal{A}(n, m)$ tal que $D^{i_1,...,i_n} = Id$, $\{i_1, ..., i_n\}$ tienen que cumplir que $menor(A; i_1, ..., i_n) \neq 0$. Según lo visto en la proposición 2.12, D existe y es única. Tenemos que $\phi(D) = \phi(A) = \phi(B)$.
Por un lado tenemos que:
$\exists C_A \in GL_n(\mathbb{K}) : C_A(A^{i_1,...,i_n}) = Id \Rightarrow \phi(C_A A) = \phi(A) = \phi(D) \Rightarrow C_A A = D$
Por otro lado:
$\exists C_B \in GL(n) : C_B(B^{i_1,...,i_n}) = Id \Rightarrow \phi(C_B B) = \phi(B) = \phi(D) \Rightarrow C_B B = D$
Entonces:
$C_A A = C_B B \Rightarrow A = C_A^{-1} C_B B$ Por tanto $A \sim B$ □

Ejemplo 2.15. *Veamos un ejemplo 2×4 de la proposición anterior. Sean A, $B \in \mathcal{A}(2,4)$:*

$$A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} 2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Tales que $\phi(A) = \phi(B) = [0:1:2:0:0:1]$. Entonces:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

*Resolviendo el sistema obtenemos **C**:*

$$C = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$$

Por lo tanto, $A \sim B$.

Ejemplo 2.16. *Ejemplo: Sean ahora $B \in \mathcal{A}(2,4)$ y $C \in GL_2(\mathbb{K})$:*

$$B = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad ; \quad C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Calculamos $A \in \mathcal{A}(2,4)$ tal que $A = CB$. Se obtiene que:

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 2 & 0 & 3 & 4 \end{pmatrix}$$

y sus menores son:

$$\phi(B) = [0:1:2:0:0:1]$$
$$\phi(A) = [0:-2:-4:0:0:-2]$$

Que definen el mismo punto proyectivo en $\mathbb{P}^{\binom{4}{2}-1}$.

Definición 2.17. *Fijemos un subconjunto $I \subset [m]$ de tamaño $n-1$ y un subconjunto $J \subset [m]$ de tamaño $n+1$. Para cada $j \in J$, definimos el signo $sgn(j, I, J)$ como $(-1)^l$, donde l es el número de elementos $i \in I$ con $i > j$ más el número de elementos $j' \in J$ con $j < j'$.*
La relación de Plücker $\mathcal{P}_{I,J}$ es la cuádrica homogénea:

$$\mathcal{P}_{I,J} = \sum_{j \in J} sgn(j; I, J) \cdot p_{I \cup j} \cdot p_{J \setminus j}$$

Donde $p_{I \cup j} = 0$ si $j \in I$.

Notemos que $\mathcal{P}_{I,J}$ es distinto de cero sólo si $|J\backslash I| \geq 3$. Si $|J\backslash I| = 3$ entonces, con unos reordenamientos adecuados, podemos escribir $I = I' \cup \{i\}$ y $J = I' \cup \{j,k,l\}$, con $i < j < k < l$. Esto implica que:

$$\mathcal{P}_{I,J} = p_{I'ij} \cdot p_{I'kl} - p_{I'ik} \cdot p_{I'jl} + p_{I'il} \cdot p_{I'jk}$$

Esta expresión que hemos obtenido se llama relación de Plücker de 3 términos.

Proposición 2.18. *Un punto de $\mathbb{P}^{\binom{m}{n}-1}$ pertenece a la Grassmanniana $Gr(n,m)$ si y sólo si es cero de todas las ecuaciones de Plücker. Por otro lado, si un punto de $\mathbb{P}^{\binom{m}{n}-1}$ tiene todas sus coordenadas distintas de cero, entonces no hace falta que se anule en todas las ecuaciones para estar en $Gr(n,m)$, sólo en las relaciones de Plücker de 3 términos.*

Demostración. Véase [5]. □

Sea $A \in \mathcal{A}(n,m)$. Sea $\hat{A}$ la matriz A con una fila extra formada por las variables x_i , $i \in \{1,2,...,m\}$. Podemos calcular las ecuaciones implícitas del espacio lineal generado por las columnas de A desarrollando por la fila de variables de los $\binom{m}{n}$ menores de la matriz $\hat{A}$.

Sea $A \in \mathcal{A}(n,m)$. Añadamos una fila extra de variables x_i , $i \in \{1,2,...,m\}$ a nuestra matriz y llamémosla $\hat{A}$. Como el rango de A es n, cualquier menor de rango $(n+1)\times(n+1)$ es igual a 0. Si calculamos todos los menores, obtendremos $\binom{m}{n}$ ecuaciones homogéneas. Éstas ecuaciones son las ecuaciones implícitas del espacio lineal generado por las columnas de A.

Ejemplo 2.19. *Sea $A \in \mathcal{A}(3,6)$. Insertemos en esta matriz una fila extra de variables, quedando $\hat{A}$ del siguiente modo:*

$$\hat{A} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Como el rango de A es 3, cualquier menor 4×4 es igual a cero. Para cada menor, desarrollamos por la primera fila obteniendo una ecuación homogénea. Por ejemplo, si cojemos las 4 primeras columnas:

$$x_1p_{234} - x_2p_{134} + x_3p_{124} - x_4p_{123} = 0 \Longrightarrow 2x_1 - x_2 = 0$$

De manera análoga, podemos obtener $\binom{6}{4} = 15$ *ecuaciones, una para cada menor* 4×4*. Son las ecuaciones implícitas del espacio lineal generado por las columnas de* A*.*
En este caso no necesitamos todas las ecuaciones para definir el m*-plano, sólo* $m - n = 6 - 3$ *(la codimensión del* n*-plano).*

No necesitamos conocer la matriz de manera explícita para obtener éstas ecuaciones. Con las coordenadas de Plücker también podemos obtenerlas. Sea $Gr(n, m)$*. Sea* $I \subseteq [m]$ *tal que* $|I| = n+1$*. Para cada* $I = \{i_1, ..., i_{n+1}\}$*, tenemos una ecuación implícita:*

$$p_{I\setminus\{i_1\}}x_{i_1} - p_{I\setminus\{i_2\}}x_{i_2} + ... \pm p_{I\setminus\{i_n\}}x_{i_n} = 0$$

En nuestro caso, las coordenadas de Plücker las tenemos calculadas del Ejemplo 2.11:

$$\phi(A) = [0 : 0 : 0 : 0 : 1 : 1 : 1 : 1 : 0 : -1 : 2 : 2 : 2 : 2 : 0 : -2 : 0 : 1 : 1 : 1]$$

De esta manera, si por ejemplo, elegimos $I = \{2356\}$*, obtenemos la siguiente ecuación:*

$$x_2p_{356} - x_3p_{256} + x_5p_{236} + x_6p_{235} = 0 \Longrightarrow x_2 + 2x_3 + 2x_5 - 2x_6 = 0$$

Esta manera de conseguir ecuaciones ímplicitas la utilizaremos más adelante para definir espacio lineal tropical.

Capítulo 3

Matroides

3.1. Definición de Matroide

A grandes rasgos, una matroide es una estructura que generaliza el concepto de independencia lineal en espacios vectoriales. Existen muchas definiciones equivalentes de matroides. Daremos algunas definiciones equivalentes de matroide, aunque hay muchas más. Primero recordemos el concepto que queremos generalizar:

Definición 3.1. *Dado un conjunto finito de vectores $v_1, v_2, ..., v_n$, se dice que son linealmente independientes si existen números $a_1, a_2, ..., a_n$, donde la ecuación:*

$$a_1 v_1 + a_2 v_2 + ... + a_n v_n = 0$$

se satisface únicamente cuando $a_1, a_2, ..., a_n$ son todos cero. En caso contrario, se dice que son linealmente dependientes.

Definición 3.2 (Matroide por conjuntos independientes)**.** *Una matroide finita M es un par $(E, \mathcal{I})$, donde E es un conjunto finito que podemos identificar con $\{1, 2, ..., m\}$ e $\mathcal{I}$ una colección de subconjuntos de E. Los elementos de $\mathcal{I}$ se llaman sunconjuntos independientes de M. Los subconjuntos de E que no pertenecen a $\mathcal{I}$ se llaman dependientes. $\mathcal{I}$ está contenido en $\mathcal{P}(E)$ y cumple las siguientes condiciones:*

- *$\mathcal{I}(1)$: $\emptyset \in \mathcal{I}$, (el vacío es independiente)*
  Además, cualquier subconjunto de un conjunto independiente es independiente, es decir, I es hereditaria.

- $\mathcal{I}(2)$*: Si* $X \in \mathcal{I}$ *y* $X' \subseteq X$ *entonces* $X' \in \mathcal{I}$.
- $\mathcal{I}(3)$*: Si* $X, Y \in \mathcal{I}$ *y* $|X| < |Y|$*, entonces existe un elemento* $x \in Y \backslash X$ *tal que* $X \cup \{x\} \in \mathcal{I}$.

Ejemplo 3.3. *Sea la matriz*

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.1}$$

Aquí hablamos de las columnas como conjuntos independientes, algunos de ellos son, por ejemplo:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad ; \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad ; \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Aunque hay muchos más. Por ejemplo:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 2 \end{pmatrix}$$

Si es un conjunto dependiente.

Teorema 3.4. *Sea* A *una matriz tal que los elementos de* E *son sus columnas y los elementos de* I *sus conjuntos linealmente independientes.* $M = (E, \mathcal{I})$ *es un matroide.*

Demostración. Es trivial que $\mathcal{I}$ cumple $\mathcal{I}(1)$ e $\mathcal{I}(2)$.
Para comprobar $\mathcal{I}(3)$:
Sean X y Y subconjuntos de E linealmente independientes tal que $|X| < |Y|$. Sea W el subespacio $\mathbb{K}^m$ generado por $X \cup Y$. Entonces, la dimensión de W, $dim(W)$, es como mínimo $|Y|$.
Ahora supongamos que $X \cup \{x\}$ es linealmente dependiente para todo $x \in \ Y \backslash X$. Entonces, W está contenido en el espacio generado de X.Así, llegamos a una contradicción:

$$|Y| \leq dim(W) \leq \ |X| \ < |Y|$$

Por lo tanto, concluimos que $Y \backslash X$ contiene un elemento x tal que $X \cup x \in \mathcal{I}$, esto es, se cumple $\mathcal{I}(3)$. □

Ejemplo 3.5. *Sea $A \in \mathcal{A}(3,4)$:*

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \end{pmatrix}$$

Entonces $M = (E, \mathcal{I})$, con:

$$E = \{1,2,3,4\} \quad ; \quad \mathcal{I} = \{1,2,3,4,13,14,23,24,34,124,234\}$$

Es una matroide.

Definición 3.6. *Un circuito de una matroide M es un conjunto dependiente minimal.*

Definición 3.7. *Decimos que una matroide es realizable cuando la matriz del Teorema 3.4 existe. En caso contrario se dice que la matroide es no realizable.*

Veamos otra definición equivalente de matroide:

Definición 3.8 (Matroide por circuitos)**.** *Una matroide M es un par (E, C), donde E es un conjunto finito que podemos identificar con $1, 2, ..., m$ y C son una serie de subconjuntos no vacíos de E llamados circuitos, que cumplen las siguientes propiedades:*

- *Ningún conjunto propio de un circuito es un circuito.*
- *Si c_1, c_2 son circuitos distintos y $e \in c_1 \cap c_2$, entonces $(c_1 \cup c_2) \backslash e$ contiene un circuito.*

Notemos que un conjunto A será dependiente si existe un circuito C con $C \subseteq A$ e independiente si no contiene circuitos.

Ejemplo 3.9. *Sea la matriz del Ejemplo 3.3. Sus circuitos son:*

$$C = \{\{1,2\},\{3,4,5\},\{1,4,6\},\{2,4,6\},\{1,3,5,6\},\{2,3,5,6\}\}$$

Donde cada número indica la posición de la columna de la matriz A referida.

Veamos ahora un ejemplo de matroide no realizable:

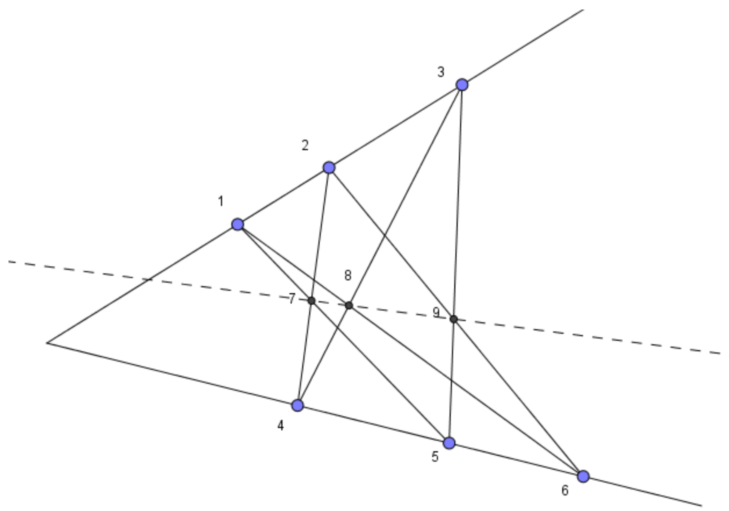


Figura 3.1: Representación gráfica del Teorema de Pappus

Ejemplo 3.10. *Sea $M = (E, C)$ con $E = \{1, 2, ..., 9\}$ y $C = \{123, 157, 168, 269, 247, 348, 359, 456\} \cup \{$todo subconjunto de tamaño 4 que no contenga los circuitos de tamaño $3\}$. Se puede comprobar que esto es una matroide.*
Esta matroide no es realizable sobre $\mathbb{C}$ porque estos circuitos se corresponden con las hipótesis del Teorema de Pappus [12], véase la Figura 3.1. El Teorema afirma que entonces 7, 8 y 9 están alineados. Esto implica que cualquier matroide realizable que tenga estos circuitos tiene que tener también al circuito 789.

Uno de los objetos más básicos y fundamentales en el álgebra lineal son las bases. A continuación veremos que también podemos definir una matriz por sus bases.

Definición 3.11. *Una base es un conjunto independiente maximal.*

Definición 3.12 (Matroide por bases)**.** *Una matroide M es un par (E, β), donde E es un conjunto finito que podemos identificar con $\{1, 2, ..., m\}$ y β una serie de subconjuntos no vacíos llamados bases de E que cumplen la siguiente propiedad conocida como el teorema del intercambio:*
Si σ y σ' son bases y se tiene que dado $i \in \sigma \backslash \sigma'$, entonces existe un elemento $j \in \sigma' \backslash \sigma$ tal que $(\sigma \backslash \{i\}) \cup \{j\}$ es una base.

Esta propiedad de las bases implica la siguiente propiedad más fuerte: el elemento $j \in \sigma \backslash \sigma'$ puede escogerse, tal que $(\sigma' \backslash \{j\}) \cup \{i\}$ también es una base.

Diremos que un conjunto $X \subset \beta$ es independiente si existe una base $B \in \beta$ con $X \subset B$.

Ejemplo 3.13. *Sea la matriz del Ejemplo 3.3. Sus bases son:*

$$\beta = \{(1,3,4),(1,3,5),(1,3,6),(1,4,5),(1,5,6),(2,3,4),(2,3,5),$$
$$(2,3,6),(2,4,5),(2,5,6),(3,4,6),(3,5,6),(4,5,6)\}$$

Ahora que tenemos definido lo que es una matroide, veamos que relación tiene con la Grassmanniana:

Proposición 3.14. *Cada punto P de $Gr(n,m)$ define una matroide. Sea $A \in \mathcal{A}(n,m)$ con $\phi(A) = P$. Defino la matroide $M = (E, \beta)$ de la siguiente manera: $E = \{1, ..., m\}$, $\{i_1, ..., i_n\} \in \beta$ si y sólo si las columnas de $i_1, ..., i_n$ son independientes. Veamos ahora la misma matroide a partir de sus coordenadas de Plücker. $M = (E, \beta)$ será la matroide con $E = \{i_1, ..., i_n\}$, $\beta = \{i_1, ..., i_n : p_{i_1,...,i_n} \neq 0\}$.*

Ya hemos visto lo que son las Grassmannianas $Gr(n,m)$ y las matroides M. Más adelante veremos que podemos agrupar todos los puntos de una Grassmanniana asociados a una misma matroide. Conseguiremos así una serie de subconjuntos de $Gr(n,m)$ que forman una partición y, a cada uno de estos subconjuntos lo llamaremos $Gr(M,n,m)$.

Capítulo 4

Geometría tropical

4.1. Introducción

La geometría tropical es una rama relativamente nueva de las matemáticas. Es curioso el adjetivo de *tropical*, pero realmente no tiene ningún significado profundo. Se trata de un homenaje que le hicieron unos matemáticos franceses al matemático e informático brasileño Imre Simon, autor de uno de los primeros artículos que llevan la palabra tropical "On semigroups of matrices over the tropical semiring" [9]. Realmente, existen artículos más antiguos que hablan sobre el tema aunque no lleven la palabra tropical.

4.1.1. Aritmética tropical

Definimos como semianillo tropical al conjunto $\mathbb{T} = \mathbb{R} \cup \{\infty\}$, junto a las operaciones suma y multiplicación definidas como:

$$x \oplus y := \text{mín}(x, y) \qquad y \qquad x \odot y := x + y$$

El término semianillo se debe a que $\mathbb{T}$ satisface los siguientes axiomas:

1. $(\mathbb{T}, \oplus, 0_{\mathbb{T}})$ es un monoide conmutativo con elemento identidad $0_{\mathbb{T}} = \infty$:

    a) $(x \oplus y) \oplus z = x \oplus (y \oplus z)$

    b) $0_{\mathbb{T}} \oplus x = x \oplus 0_{\mathbb{T}} = x$

    c) $x \oplus y = y \oplus x$

2. $(\mathbb{T}, \odot, 1_{\mathbb{T}})$ es un grupo abeliano con elemento identidad $1_{\mathbb{T}} = 0$:

a) $(x \odot y) \odot z = x \odot (y \odot z)$

b) $1_{\mathbb{T}} \odot x = x \odot 1_{\mathbb{T}}$

c) $x \odot y = y \odot x$

d) $x \odot x^{-1} = 1_{\mathbb{T}}$, si $x \neq 0_{\mathbb{T}}$

3. La multiplicación cumple con la ley distributiva con respecto a la suma:

   a) $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$

   b) $(x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$

4. $0_{\mathbb{T}} \odot x = x \odot 0_{\mathbb{T}} = 0_{\mathbb{T}}$

Existen trabajos en los que para la suma se toma el máximo en lugar del mínimo.
Es importante notar que en este anillo la suma es idempotente, ya que:

$$x \oplus x = x \quad \forall x \in \mathbb{T}$$

Esto simplifica mucho las operaciones, por ejemplo:

$$(x \oplus y)^{\odot n} = (x)^{\odot n} \oplus (y)^{\odot n}$$

(donde el exponente indica el producto tropical n veces)
A partir de ahora, denotaremos $x^{\odot n}$ como x^n.

Nota 4.1. *La suma no es cancelativa:*

$$a \oplus b = a \oplus c \not\Rightarrow b = c$$

Ejemplo 4.2.

$$1 \oplus 3 = 1 \oplus 7 \not\Rightarrow 3 = 7$$

Proposición 4.3. *Para cualquier $n \in \mathbb{Z}$, se cumple la siguiente igualdad:*

$$(x \oplus y)^n = x^n \oplus y^n$$

Demostración.

$$(x \oplus y)^n = n*(x \oplus y) = n*(min(x,y)) = min(n*x, n*y) = min(x^n, y^n) = x^n \oplus y^n$$

□

4.2. Geometría Tropical

4.2.1. Polinomios tropicales

Sean $w_1, w_2, ..., w_n$ variables sobre en el semianillo tropical $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$. Un monomio es un producto cualquiera de estas variables, donde las repeticiones están permitidas. Por la conmutatividad podemos ordenar el producto y escribir los monomios con la notación usual con las variables elevadas a exponentes:

$$w_2 \odot w_4 \odot w_2 \odot w_1 \odot w_4 \odot w_2 = w_1 \odot w_2^3 \odot w_4^2$$

Observemos que la evaluación de un monomio tropical de n variables es una función lineal afín definida de $\mathbb{T}^n$ en $\mathbb{T}$. Por ejemplo:

$$a \odot w_1 \odot w_2^3 \odot w_4^2 = a + w_1 + 3w_2 + 2w_4 = a + \langle w, i \rangle$$

donde $w = (w_1, ..., w_4)$ es el vector de las variables, e $i = (1, 3, 0, 2) \in \mathbb{N}^4$ es el vector de sus respectivas potencias.
Por simplificar, escribiremos:

$$a \odot w_1^{i_1} \odot ... \odot w_n^{i_n} = a \odot w^{\odot i}$$

Antes de definir lo que es un polinomio tropical, introduzcamos el siguiente concepto.

Definición 4.4. *Sea $\triangle \in \mathbb{R}^n$. Diremos que $\triangle$ es un politopo entero si $\triangle$ es el envolvente convexa de una cantidad finita de puntos con coordenadas enteras.*

Definamos ahora polinomio tropical:

Definición 4.5. *Un polinomio tropical f es una combinación finita de monomios tropicales:*

$$f(w_1, ..., w_n) = \bigoplus_{i \in \mathtt{A}} a_i \odot w^i$$

donde los coeficientes $a_i \in \mathbb{T}$ para todo $i \in \mathtt{A}$, y $\mathtt{A} \subset \mathbb{Z}^n$ un conjunto finito (son polinomios con exponentes también negativos que se llaman polinomios de Laurent).

Si $a_i \in \mathbb{R}\ \forall i \in \mathtt{A}$, decimos que $\mathtt{A}$ es el soporte del polinomio tropical f. Todo polinomio tropical representa una función $\mathbb{T}^n \longrightarrow \mathbb{T}$. Cuando evaluamos estas

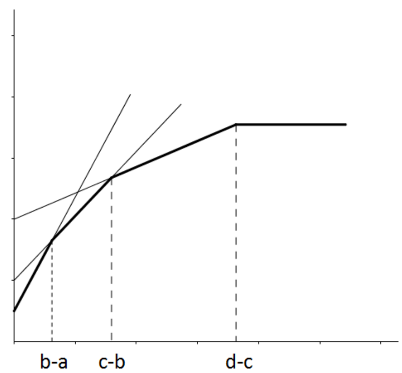


Figura 4.1: Gráfica de la función definida por el polinomio tropical $p(w) = a \odot w^3 \oplus b \odot w^2 \oplus c \odot w \oplus d$ del ejemplo (4.6)

funciones en la aritmética clásica, obtenemos el mínimo de una colección finita de funciones lineales:

$$\hat{f}(w_1, ..., w_n) = min_{i \in \mathtt{A}}\{a_i + \langle w, i \rangle\}$$

Si restringimos el dominio de esta función a $\mathbb{R}^n$ nos queda una función $p : \mathbb{R}^n \longrightarrow \mathbb{R}$ que es continua y lineal a trozos y para cada trozo lineal, la diferencial es un vector entero.

Ejemplo 4.6. *Consideremos un polinomio genérico de grado tres en una variable w_1:*

$$f(w_1) = a \odot w_1^3 \oplus b \odot w_1^2 \oplus c \odot w_1 \oplus d$$

Para hacer la gráfica de la función $\hat{f}$ que define, dibujamos cuatro líneas en el plano de coordenadas $(w_1, w_2) : w_2 = 3w_1 + a, w_2 = 2w_1 + b, w_2 = w_1 + c$ y la línea horizontal $w_2 = d$. El valor de $\hat{f}(w)$ es el mínimo de w_2 tal que (w_1, w_2) pertenece a alguna de estas líneas. Las cuatro líneas contribuyen al grafo si

$$b - a \leq c - a \leq d - c$$

En este caso, estos tres valores de w son exactamente los puntos en los cuales $\hat{f}$ no es diferenciable.

Nota 4.7. *Polinomios distintos pueden representar la misma función.*

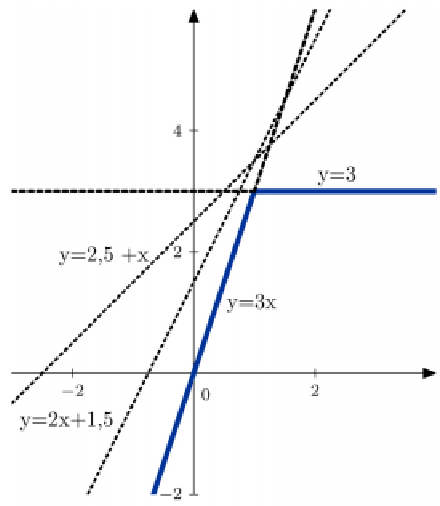


Figura 4.2: Gráfica de la función definida por el polinomio tropical del Ejemplo 4.8. Las rectas punteadas corresponden a monomios que no aportan al mínimo [2].

Ejemplo 4.8. *Consideremos un caso particular del ejemplo anterior, tomando* $p = 0 \odot w^3 \oplus 3$*. Podemos ver el gráfico de la función que define este polinomio tropical en la Figura 4.2. Si sumamos* p *a cualquier monomio del tipo* $a \odot w^2$ *con* $a > 1$*, o del tipo* $b \odot w$ *con* $b > 2$*, la función que definen será la misma, pues estos monomios se corresponden con rectas por encima del gráfico de* $\hat{p}$*, con lo que no aportan al mínimo que se alcanza en los monomios de* $\hat{p}$*.*

4.2.2. Raíces de un polinomio tropical

Dado un polinomio tropical f queremos definir sus raíces. Consideramos el polinomio en una variable:

$$f(w) := a \odot w \oplus b$$

y observamos que la ecuación $a \odot w \oplus b = 0_{\mathbb{T}}$ no tiene solución si $b \neq 0_{\mathbb{T}}$. Entonces tenemos que buscar otra definición de los ceros de f. La igualdad:

$$a \odot w \oplus b = a \odot (w \oplus (b - a))$$

sugiere definir, así como en el caso clásico, el número $b - a$ como cero de f.

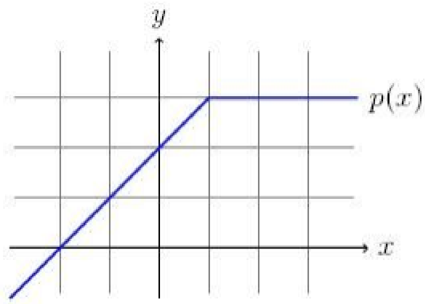


Figura 4.3: Gráfica de la función definida por el polinomio tropical del Ejemplo 4.9 [2].

Ejemplo 4.9. *Sea* $f := 2 \odot w \oplus 3$. *De acuerdo con la definición precedente,* $w = 1$ *es cero de* f. *La gráfica de* $f(w)$ *es singular para este de valor de* w *puesto que en ese punto, el mínimo entre* $2 + w$ *y 3 se alcanza dos veces.*

Esto nos motiva a dar la siguiente definición:

Definición 4.10. *Llamamos conjunto de ceros de un polinomio tropical* $f = \bigoplus_{i \in \mathtt{A}} c_i \odot w^i$ *al conjunto* $Z(f)$ *de todos los puntos de* $w \in \mathbb{R}$ *donde el valor* $f(w) = min_{i \in \mathtt{A}}\{c_i + \langle w, i \rangle\}$ *se alcanza para al menos dos índices distintos* $i, j \in \mathtt{A}$ *o, equivalentemente, donde* $\hat{f}$ *no es diferenciable.*

4.2.3. Módulo tropical

Una posible definición de módulo tropical podría ser la siguiente:

Definición 4.11. *Sea* V *un conjunto con las siguientes dos operaciones:*

$$\oplus : V \times V \longrightarrow V \quad \odot : \mathbb{T} \times V \longrightarrow V$$

que satisface las siguientes propiedades:

- V *es un monoide con respecto a* $\oplus$.
- $a \odot v \in V$ *para todo* $a \in \mathbb{T}$ *y* $v \in V$.
- $(a \odot b) \odot v = a \odot (b \odot v)$.

- $(a \oplus b) \odot v = (a \odot v) \oplus (b \odot v)$, *para todo* $a, b \in \mathbb{T}$ *y* $v \in V$.
- $a \odot (v_1 \oplus v_2) = (a \odot v_1) \oplus (a \odot v_2)$, *para todo* $a \in \mathbb{T}$ *y* $v_1, v_2 \in V$.
- $0 \odot v = v$, *para todo* $v \in V$.

Ejemplo 4.12. *Sea* $V = \mathbb{T}^2$ *un* $\mathbb{T}$*-módulo generado por los vectores* $v_1 = (1, 0)$ *y* $v_2 = (0, 1)$. *Entonces existe un vector* $v = (x, y)$ *tal que,*

$$\begin{pmatrix} x \\ y \end{pmatrix} = a \odot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus b \odot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

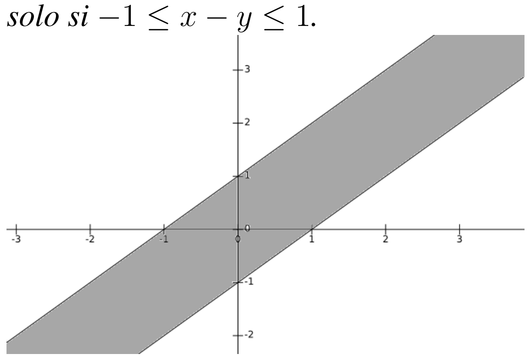
Que es el siguiente sistema de ecuaciones:

$$\begin{cases} x = a \odot 1 \oplus b \odot 0 \\ y = a \odot 0 \oplus b \odot 1 \end{cases}$$

por lo tanto,

$$\begin{cases} x = min\{a + 1, b + 0\} \\ y = min\{a + 0, b + 1\} \end{cases}$$

Y tiene solución si y solo si $-1 \leq x - y \leq 1$.



La zona sombreada es la zona en la que existe solución.

En este ejemplo hemos comprobado que esta definición de módulo tropical no es “buena”. En el siguiente capítulo, vamos a utilizar la Grassmanniana para definirlo.

Capítulo 5

Dressiano Tropical

El objetivo de este capítulo es definir un análogo tropical de la Grassmanniana clásica $Gr(n,m)$.

Como vimos en el primer capítulo de este trabajo, $Gr(n,m)$ parametriza subespacios n lineales de un espacio m dimensional. Cada punto de una Grassmanniana puede definir diferentes espacios lineales. Recordémoslo con el siguiente ejemplo:

Ejemplo 5.1. *Sea $Gr(3,5)$ y los espacios asociados a las siguientes matrices:*

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 5 \\ 3 & 0 & 2 & 0 & 1 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} 2 & 0 & 3 & 1 & -1 \\ 0 & 0 & 2 & 5 & 0 \\ 1 & 0 & -4 & 2 & 0 \end{pmatrix}$$

Sus coordenadas de Plücker son:

$$\phi(A) = [11 : 0 : 25 : 0 : -18 : 0 : 0 : -23 : 0 : 0]$$

$$\phi(B) = [0 : 0 : 0 : -11 : 2 : 5 : 0 : 0 : 0 : 24]$$

A define la matroide $M_1 = (E, \beta_1)$, definida por bases:

$$E = \{1,2,3,4,5\} \quad ; \quad \beta_1 = \{\{123\},\{125\},\{135\},\{235\}\}$$

B define la matroide $M_2 = (E, \beta_2)$:

$$E = \{1,2,3,4,5\} \quad ; \quad \beta_2 = \{\{134\},\{135\},\{145\},\{345\}\}$$

Las matroides de estos dos espacios son diferentes, por lo tanto, no son el mismo.

A continuación introduciremos $Gr(M,n,m)$ una partición de $Gr(n,m)$ con el que agruparemos todos los puntos asociados a la misma matroide.

Definición 5.2. *Sea $M=(E,\beta)$ una matroide con bases β. Definimos $Gr(M,n,m)$ como el subconjunto de $Gr(n,m)$ donde cada punto define a la matroide M:*

$$Gr(M,n,m)=Gr(n,m)\cap\{p_A=0:A\notin\beta\}\cap\{p_A\neq 0:A\in\beta\}$$

La matroide M indica los p_A que son o no iguales a cero. $Gr(M,n,m)$ es vacío si M es no realizable sobre $\mathbb{K}$.

Ejemplo 5.3. *Sea $A\in\mathcal{A}(2,4)$ la siguiente matriz:*

$$A=\begin{pmatrix}0 & 2 & 3 & 0\\ 1 & -1 & 0 & 2\end{pmatrix}$$

Tenemos $\phi(A)$ y la matroide $M=(E,\beta)$ asociada a A:

$$\phi(A)=[-2:-3:0:3:3:3]$$

$$E=\{1,2,3,4\}\quad;\quad\beta=\{\{12\},\{13\},\{23\},\{24\},\{34\}\}$$

Calculo las ecuaciones de Plücker. Si $I=\{1\}$ y $J=\{2,3,4\}$ obtenemos:

$$P_{I,J}=p_{12}p_{34}\pm p_{13}p_{24}\pm p_{14}p_{23}$$

Podemos escoger I de otras 3 maneras posibles, pero obtenemos la misma ecuación. Entonces, los puntos que pertenecen a $Gr(M,n,m)$ son aquellos que cumplen $P_{I,J}$ con $p_{12}\neq 0$, $p_{13}\neq 0$, $p_{14}=0$, $p_{23}\neq 0$, $p_{24}\neq 0$, $p_{34}\neq 0$.

Ahora introduciremos el Dressiano, un concepto puramente tropical. Los desarrolladores de esta teoría decidieron usarlo porque posee buenas propiedades geométricas.

Definición 5.4. *Sean $\mathcal{P}_{I,J}$ las relaciones cuadráticas de Plücker introducidas en la Definición 2.17:*

$$\mathcal{P}_{I,J}=\sum_{j\in J}sgn(j;I,J)\cdot p_{I\cup j}\cdot p_{J\setminus j}$$

Las ecuaciones tropicales de Plücker son su tropicalización directa:

$$\bigoplus_{j\in J}0\odot p_{I\cup j}\odot p_{J\setminus j}$$

Definición 5.5. *El Dressiano $Dr(M,n,m)$ es un subconjunto del semianillo tropical $\mathbb{T}^{\binom{m}{n}}$. Un punto pertenece $Dr(M,n,m)$ si satisface las ecuaciones de Plücker tropicales (podemos quedarnos con los 3-terms) y cada coordenada de $\mathbb{T}^{\binom{m}{n}}$, indexada por un subconjunto de n elementos de m, que van a ser ∞ si el subconjunto no es base de M y un número disntinto de ∞ en caso de que si lo sea. Es decir:*

$$Dr(M,n,m) = (\bigcap_{I,J} Z(P_{I,J}^{trop})) \cap \{p_A \neq \infty : A \in \beta\} \cap \{p_A = \infty : A \notin \beta\}$$

Observación 5.6. *Cada punto de una Grassmanianna nos da un espacio lineal. El Dressiano es lo más parecido al análogo tropical, ya que cada punto nos da un espacio lineal tropical.*

Ejemplo 5.7. *Calculemos los puntos del Dressiano dados A, M, n, m los del Ejemplo 5.3. Al tropicalizar la ecuación de Plücker obtenemos:*

$$P_{I,J}^{trop} = p_{12} \odot p_{34} \oplus p_{13} \odot p_{24} \oplus p_{14} \odot p_{23}$$

El neutro para $\oplus$ es ∞, por lo que $p_A = \infty$ si $A \notin \beta$ y $p_A \neq \infty$ si $A \in \beta$.

Entonces, los puntos que pertenecen a $Dr(M,n,m)$ son aquellos que cumplen $P_{I,J}^{trop}$ y $p_{12} \neq \infty$, $p_{13} \neq \infty$, $p_{14} = \infty$, $p_{23} \neq \infty$, $p_{24} \neq \infty$, $p_{34} \neq \infty$.

Algunos ejemplos de puntos que están en el Dressiano son:

$$[3 : 6 : \infty : 901 : -1 : 2]$$

$$[-5 : -1 : \infty : 6 : 0 : 4]$$

$$[7 : 7 : \infty : 7 : 7 : 7]$$

Teorema 5.8. *Sea M una matroide. El Dressiano $Dr(M,n,m)$ nunca es vacío.*

Se puede probar que el punto con $p_{i_1,\ldots,i_n}$ igual a 0 si $i_1, \ldots, i_n$ es una base de M e ∞ en caso de que no lo sea siempre pertenece al Dressiano $Dr(M,n,m)$.

Ejemplo 5.9. *Sea $Dr(M,3,5)$ con $M = (E,\beta)$:*

$$M = (\{1,2,3,4,5\}, \{\{123\}, \{124\}, \{125\}, \{135\}, \{234\}, \{235\}, \{245\}, \{345\}\})$$

Sus coordenadas de Plücker tropicales son las siguientes:

$$\phi = [0 : 0 : 0 : \infty : 0 : \infty : 0 : 0 : 0 : 0]$$

Las ecuaciones de Plücker siempre van a tener solución ya que el mínimo siempre está en el 0.

Introducimos ahora un subconjunto del Dressiano, La Grassmanianna Tropical.

Definición 5.10. *La Grassmanniana Tropical $GRT(M, n, m)$ es un subconjunto de $\mathbb{T}^{\binom{m}{n}}$. Un punto pertenece a $GRT(M, n, m)$ si es solución de las ecuaciones tropicalizadas del ideal clásico generado por las ecuaciones de Plücker tropicales:*

$$\mathcal{I}(GR(M)) = (\sum f_i P)$$

Donde f_i es un polinomio cualquiera y P^i una ecuación de Plücker.
Fijémonos en:

$$\sum f_i P_i = \sum a_i p^i$$

que al tropicalizarlo nos queda:

$$\bigoplus 0 \odot p^i$$

La Grassmanniana Tropical $GRT(M, n, m)$ cumple las ecuaciones del Dressiano más todas las generadas por el ideal $\mathcal{I}(GR(M))$. Por lo tanto

$$GRT(M, n, m) \subseteq Dr(M, n, m)$$

Nota 5.11. *Para $n \neq 2$ se sabe que $Dr(M, n, m)$ y $GRT(M, n, m)$ pueden ser distintos, pero para $n = 2$ se ha demostrado que son iguales. En este caso particular, estos espacios coinciden con el espacio de árboles filogenéticos, que estudiaremos en el siguiente capítulo. Véase [5], página 172.*

Ejemplo 5.12. *Sean A, M, n, m del Ejemplo 5.3. Los puntos de $GRT(M, n, m)$ no sólo deben de cumplir la ecuación de Plücker tropicalizada del Ejemplo 5.7, deben cumplir las ecuaciones tropicalizadas del ideal clásico que generan estas:*

$$\mathcal{I}(Gr(M, n, m)) = (p_{12}p_{34} \pm p_{13}p_{24} \pm p_{14}p_{23})$$

Es decir, si por ejemplo tenemos la ecuación $2w_1^2w_3 + 3w_1w_2^2 + 8w_2w_3^2$, su homóloga tropical es $0 \odot w_1^2 \odot w_3 \oplus 0 \odot w_1 \odot w_2^2 \oplus 0 \odot y \odot w_3^2$. Los puntos que cumplan las infinitas ecuaciones que forman el ideal, son los que pertenecen a $GRT(M, n, n)$.

Comprobar que un punto pertenece al Dressiano puede hacerse. Para que un punto pertenezca a la Grassmanniana Tropical ha de cumplir infinitas ecuaciones, por lo que no puede comprobarse de manera trivial.

A final del Capítulo 1 vimos un método para calcular las ecuaciones implícitas de un espacio lineal generado por las columnas de $A \in \mathcal{A}(n,m)$. De manera análoga podemos calcular las ecuaciones implícitas de un espacio lineal tropical. Ésto nos servirá como definición.

Definición 5.13. *Sea $Dr(M,n,m)$. Sea P un punto del Dressiano. Sea $I \subseteq [m]$ tal que $|I| = n+1$. Para cada $I = \{i_1, ..., i_{n+1}\}$ definimos la siguiente ecuación:*

$$p_{I\setminus\{i_1\}} \odot x_{i_1} \oplus p_{I\setminus\{i_2\}} \odot x_{i_2} \oplus ... \oplus p_{I\setminus\{i_n\}} \odot x_{i_n}$$

Como hicimos en el Ejemplo 2.19. En total hay $\binom{m}{n}$ ecuaciones. Definimos el espacio lineal tropical de P como la solución de estas ecuaciones.

Ejemplo 5.14. *Sea $Dr(M,4,6)$. con una matroide asociada:*

$$M = (E,\beta) = (\{1,2,3,4,5,6\}, \{\{1346\},\{1456\},\{2346\},\{2456\},\{3456\}\})$$

Sus coordenadas de Plücker tropicales de M son:

$$\phi = [\infty : \infty : \infty : \infty : \infty : \infty : \infty : 0 : \infty : 0 : \infty : 0 : \infty : 0 : 0]$$

Según la definición, tenemos 6 ecuaciones que son las siguientes:
Para $\{12345\}$*:* $\infty \odot x_1 \oplus \infty \odot x_2 \oplus \infty \odot x_3 \oplus \infty \odot x_4 \oplus \infty \odot x_5$
Para $\{12346\}$*:* $0 \odot x_1 \oplus 0 \odot x_2 \oplus \infty \odot x_3 \oplus \infty \odot x_4 \oplus \infty \odot x_6$
Para $\{12356\}$*:* $\infty \odot x_1 \oplus \infty \odot x_2 \oplus \infty \odot x_3 \oplus \infty \odot x_5 \oplus \infty \odot x_6$
Para $\{12456\}$*:* $0 \odot x_1 \oplus 0 \odot x_2 \oplus \infty \odot x_4 \oplus \infty \odot x_5 \oplus \infty \odot x_6$
Para $\{13456\}$*:* $0 \odot x_1 \oplus 0 \odot x_3 \oplus \infty \odot x_4 \oplus 0 \odot x_5 \oplus \infty \odot x_6$
Para $\{23456\}$*:* $0 \odot x_2 \oplus 0 \odot x_3 \oplus \infty \odot x_4 \oplus 0 \odot x_5 \oplus \infty \odot x_6$

En resumen, son estas ecuaciones:

$$0 \odot x_1 \oplus 0 \odot x_2$$

$$0 \odot x_1 \oplus 0 \odot x_3 \oplus 0 \odot x_5$$

$$0 \odot x_2 \oplus 0 \odot x_3 \oplus 0 \odot x_5$$

Que precisamente cada una tiene como soporte uno de los circuitos de la matroide M. Esto ocurre siempre.

Capítulo 6

Árboles filogenéticos

En este capítulo analizaremos el caso $n = 2$, en el que el Dressiano y la Grassmanniana Tropical coincinden.

Teorema 6.1. *Cuando $n = 2$, se cumple: $Dr(M, 2, m) = GRT(M, 2, m)$*

Demostración. Véase [5]. $\square$

6.1. Árboles filogenéticos

Definición 6.2. *Un árbol filogenético es un árbol con m hojas etiquetadas y sin vértices de grado 2 (el número de aristas adyacentes al vértice). Estos, principalmente, surgen en biología, donde las etiquetas representan diferentes datos y la estructura de árbol registra su historia evolutiva. Las m aristas adyacentes a las hojas de un árbol τ son las aristas colgantes de τ.*

Ejemplo 6.3. *En la Figura 6.1 se muestra un ejemplo de árbol filogenético.*

Definición 6.4. *Una distancia de árbol es un vector $d = (d_{i,j}) \in \mathbb{R}^{\binom{m}{2}}$ construido como sigue. Sea τ un árbol filogenético. Asignamos una longitud $l_e \in \mathbb{R}$ a cada arista. Esta longitud en principio puede ser 0 o incluso un número negativo. Entre dos hojas $i, j \in \tau$ hay un único camino en τ. Definimos la distancia $d_{i,j}$ entre las hojas i, j como la suma de la longitud de todas las aristas l_e a lo largo de este camino. El vector resultante $d = (d_{i,j}) \in \mathbb{R}^{\binom{m}{2}}$ se llama distancia de árbol. Cuando todas las l_e son no negativas, la distancia de árbol especifica una métrica en el conjunto $[m] = 1, ..., m$ de hojas. Los espacios métricos que surgen de estos árboles métricos τ se llaman métricas de árbol.*

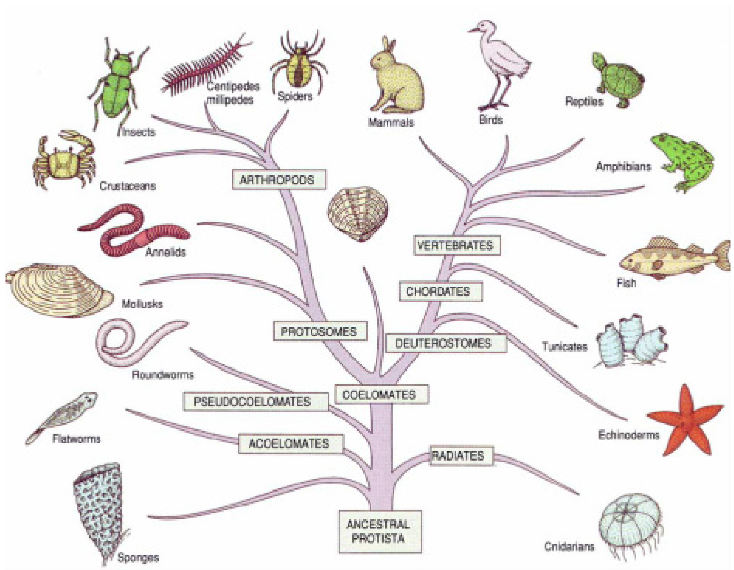


Figura 6.1: Ejemplo de árbol filogenético [4]

Definición 6.5. *Denotemos Δ al conjunto de todas las distancias de árbol en $\mathbb{R}^{\binom{m}{2}}$. Este conjunto se conoce como el espacio de los árboles filogenéticos de m hojas. Añadir el vector $\lambda \cdot \sum_{j \neq i} e_{ij}$ a la distancia de árbol d corresponde a añadir la constante $\lambda/2$ a la longitud del arista colgante i-ésimo del árbol τ. Esto demuestra que $d + (\lambda/2) \cdot \sum_{j \neq i} e_{ij} \in \Delta$. Concluimos con que el subespacio L está contenido en Δ:*

$$\Delta + L = \Delta$$

Recordemos los axiomas que definen a una métrica:

Definición 6.6. *Una métrica es un concepto que generaliza la idea geométrica de distancia. Un conjunto que tiene una métrica definida recibe el nombre de espacio métrico. Dado un conjunto $[m]$, una métrica en $[m]$ es una función:*

$$d : [m] \times [m] \to \mathbb{R}$$

que posee las siguientes propiedades:

- *Es positivamente definida, es decir: $d(x, y) \geq 0$ para todos los $x, y \in [m]$.*
- *Es simétrica: $d(x, y) = d(y, x)$ para todos los $x, y \in [m]$.*

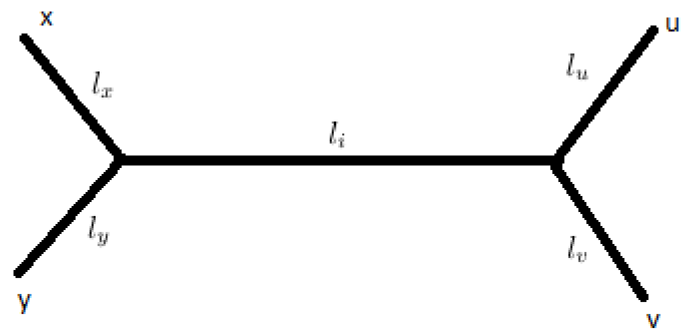


Figura 6.2: Árbol de 4 hojas

- *Cumple la desigualdad triangular, es decir, que para todos los* $x, y, x \in [m]$, *se cumple:* $d(x, z) \leq d(y, x) + d(y, z)$.
- *Es nula para puntos coincidentes:* $d(x, y) = 0 \Leftrightarrow x = y$

6.2. Condición de cuatro puntos

Definición 6.7. *A cualquier subárbol* τ *formado por cuatro hojas* u, v, x, y, *lo llamamos quartet* $(uv; xy)$ *si* u *tiene un vértice adyacente en común con* v *y* x *lo tiene con* y. *Véase la Figura 6.2.*

Teorema 6.8 (Condición de cuatro puntos)**.** *Una métrica* d *en un conjunto finito* $[m]$ *es una metrica de árbol si y solo si para cualquier cuatro elementos* $u, v, x, y \in [m]$ *el máximo de los tres números* $d_{uv} + d_{xy}, d_{ux} + d_{vy}, d_{uy} + d_{vx}$ *se alcanza al menos dos veces.*

Demostración. "$\Longrightarrow$"
Supongamos que d es igual a la métrica d_τ definida por un árbol τ en $[m]$. Entonces para cualquier quartet $(uv; xy)$ de τ, como el de la Figura 6.2, se cumple:

$$(l_u + l_v) + (l_x + l_y) \leq (l_u + l_i + l_x) + (l_v + l_i + l_y) = (l_u + l_i + l_y) + (l_v + l_i + l_x)$$

Es decir,

$$d_{uv} + d_{xy} \leq d_{ux} + d_{vy} = d_{uy} + d_{vx} \tag{6.1}$$

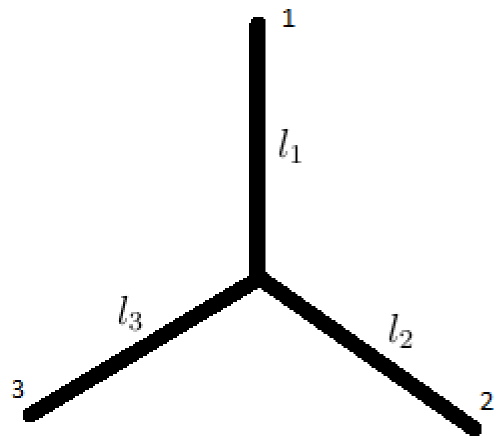


Figura 6.3: Árbol de 3 hojas

Por lo tanto se cumple que el máximo se alcance dos veces.
"$\Longleftarrow$"
La otra parte de la implicación la vamos a probar aplicando inducción en el número de hojas m. Para $m = 3$, tenemos el caso de la Figura 6.3. A partir de las distancias de árbol d_{ij}, podemos calcular las longitudes l_i con el siguiente sistema de ecuaciones:

$$d_{12} = l_1 + l_2$$
$$d_{13} = l_1 + l_3$$
$$d_{22} = l_2 + l_3$$

El sistema es compatible determinado, por lo que para $m = 3$ siempre se cumple.

Supongamos $m \geq 4$, y que el teorema es válido para todos los espacios métricos con menos de m elementos. Enconces sea d una métrica en $[m] = \{1, 2..., m\}$ que satisfaga la condición de los cuatro puntos.
Escogemos i, j, k que maximicen $d_{ik} + d_{jk} - d_{ij}$. Por hipótesis de inducción hay un árbol τ' en $[m]\backslash i$ que realiza d restringida a $[m]\backslash i$.
Sea λ la longitud de la arista colgante e de τ' adyacente a j. Subdividimos e añadiendo un nuevo vértice interior conectado a la hoja i, como vemos en la Figura 6.4.

La arista adyacente a j se le asigna una longitud $\lambda_i = (d_{ij} + d_{ik} - d_{jk})/2$, a la arista adyacente a j le asignamos una longitud de $\lambda_j = (d_{ij} + d_{jk} - d_{ik})/2$ y por

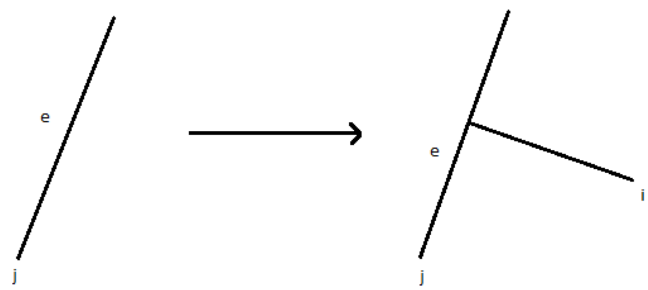


Figura 6.4: Subvidimos e con un nuevo vértice y le conectamos con i

último, a la parte restante de e le asignamos longitud $\lambda - \lambda_j$.
Veamos que el árbol resultante τ no tiene ninguna arista con peso negativo, y se cumple que $d = d_\tau$. Por construcción, d y d_τ coinciden en todos los pares $x, y \in [m]\backslash i$.
Sea l cualquier hoja de τ' diferente de i, j, k. Nuestra elección de i, j, k implica que $d_{ik} + d_{jk} - d_{ij} \geq d_{kl} + d_{ik} - d_{il}$ y $d_{ik} + d_{jk} - d_{ij} \geq d_{kl} + d_{jk} - d_{jl}$.
La condición del cuarto punto de nuestra hipótesis da entonces

$$d_{ij} + d_{kl} \leq d_{ik} + d_{jl} = d_{il} + d_{jk}$$

Como d es una métrica, tenemos que $\lambda_i \geq 0$ y $\lambda_j \geq 0$. Veamos ahora que $\lambda - \lambda_j$ es no negativo, fijemos una hoja $l \neq j, k$ de τ'. Entonces se tiene que $\lambda = (d_{jk} + d_{jl} - d_{kl})/2$, asi que

$$\lambda - \lambda_j = (d_{ik} + d_{jl} - d_{ij} - d_{kl})/2 \geq 0$$

Por lo tanto, nuestro árbol τ no tiene aristas con pesos negativos, pero puede ser 0. Además tenemos que $(d_\tau)_{ij} = \lambda_i + \lambda j = d_{ij}$ y que $(d_\tau)_{il} = (d_\tau)_{jl} - \lambda_j + \lambda_i = d_{jl} + d_{ik} - d_{jk} = d_{il}$ para $l \neq i$. $\square$

Ejemplo 6.9. *Sea el árbol métrico de 4 hojas de la Figura 6.5.*
Tomando $p_{i,j} = -d(i, j)$, veamos que se cumple:

$$P_{I,J}^{trop} = p_{12} \odot p_{34} \oplus p_{13} \odot p_{24} \oplus p_{14} \odot p_{23}$$

$$\begin{aligned} P_{I,J}^{trop} &= min\{-((1+2)+(3+4)), -((1+5+3)+(2+5+4)), \\ &- ((1+5+4)+(2+5+3))\} = min\{-10, -20, -20\} = -20 \end{aligned}$$

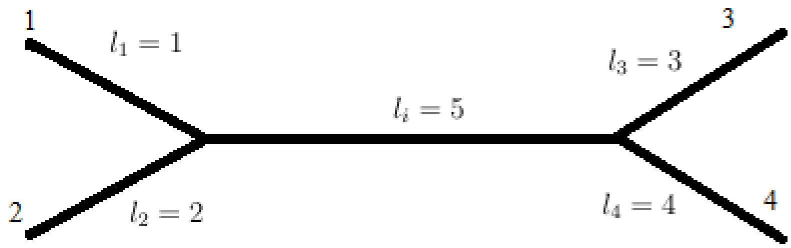


Figura 6.5: Ejemplo de árbol

Como vemos, la condición de cuarto puntos no son más que las ecuaciones de Plücker cambiadas de signo. Es decir, los árboles filogenéticos son puntos del Dressiano.

Teorema 6.10. *El conjunto de métricas de árbol de m hojas es igual al subconjunto de $-Dr(M, 2, m)$ con todas las coordenadas positivas, donde $M = (E, \beta)$ es la matroide uniforme (todos los subconjuntos de E de 2 elementos son bases).*

6.3. Caso real de árbol filogenético

En esta sección intentaremos recrear un árbol filogenético a partir del ADN mitocondrial de los siguientes humanos: Hemos ido a una base de datos de ADN mitocondrial humano [14] y hemos descargado las muestras que tienen el siguientes identificador:

AF381982, AY195748, AF346989, AF346963, AF346988.

Que vamos a llamar, 1, 2, 3, 4 y 5 respectivamente.

La distancia que vamos a usar es el número de nucleótidos diferentes en cada especie. Por ejemplo la distancia entre:

AGTCA y GATCA

es 2 porque difieren en la primera y segunda posición. El problema es que el ADN no está alineado.

Por ejemplo, entre las posiciones 515 y 540 hemos obtenido lo siguiente:

Cuadro 6.1: Distancias entre las 5 hojas del árbol

	1	2	3	4	5
1	0	41	47	34	28
2		0	18	39	41
3			0	45	47
4				0	32
5					0

$$\begin{array}{ll} 1: & GCACACACACACCGCTGCTAACCCC \\ 2: & GCACACACACACCGCTGCTAACCCC \\ 3: & GCACACACACCGCTGCTAACCCCAT \\ 4: & GCACACACACACCGCTGCTAACCCC \\ 5: & GCACACACACACCGCTGCTAACCCC \end{array}$$

Vemos que la 3 está muy alejada de las demás, pero más que eso, lo que parece es que han desaparecido 2 nucleótidos. Añadimos una nueva letra 'X' para indicar que un nucleótido ha desaparecido. Hemos añadido 2 'X' y las secuencias quedan de este modo:

$$\begin{array}{ll} 1: & GCACACACACACCGCTGCTAACCCCAT \\ 2: & GCACACACACACCGCTGCTAACCCCAT \\ 3: & GCACACACA\mathbf{XX}CCGCTGCTAACCCCAT \\ 4: & GCACACACACACCGCTGCTAACCCCAT \\ 5: & GCACACACACACCGCTGCTAACCCCAT \end{array}$$

En este ejemplo lo hemos hecho a mano, pero un problema interesante es buscar definir un algoritmo que decida cómo alinear ADN.

Una vez alineado, comparando las diferencias entre los ADN, hemos obtenido una serie de distancias, reflejadas en el Cuadro 6.1, donde la distancia ha sido calculada como el número de nucleótidos diferentes entre cada secuencia.

De esta tabla, sacamos 5 condiciones de cuatro puntos:

Con las hojas $\{1,2,3,4\}$, $\{1,2,3,5\}$ y $\{2,3,4,5\}$, la condición de los cuatro puntos se cumple:

$$\begin{array}{lll} d_{12}+d_{34}=86 & d_{12}+d_{35}=88 & d_{23}+d_{45}=50 \\ d_{13}+d_{24}=86 & d_{13}+d_{25}=88 & d_{24}+d_{35}=86 \\ d_{14}+d_{23}=52 & d_{15}+d_{23}=46 & d_{15}+d_{43}=86 \end{array}$$

Cuadro 6.2: Distancias finales entre las 5 hojas del árbol

	1	2	3	4	5
1	0	41	47	34	28
2		0	18	39	41
3			0	45	47
4				0	**34**
5					0

Sin embargo, observamos que escogiendo las hojas $\{1,3,4,5\}$ y $\{1,2,4,5\}$ la condición de los cuatro puntos no se cumple:

$$
\begin{aligned}
d_{13} + d_{45} &= 79 \qquad & d_{12} + d_{45} &= 73 \\
d_{14} + d_{35} &= 81 \qquad & d_{14} + d_{25} &= 75 \\
d_{15} + d_{43} &= 73 \qquad & d_{15} + d_{24} &= 67
\end{aligned}
$$

Eso quiere decir que no tenemos una métrica de árbol. Para poder construir un árbol usando el Teorema 6.8, tenemos que hacer un pequeño ajuste a los datos. Un posible cambio que vemos a simple vista es aumentar en 2 unidades a d_{45}, obteniendo el Cuadro 6.2. Aquí lo hemos hecho a ojo, pero un problema interesante fuera del alcance de este trabajo, es dada una métrica que no es de árbol, ajustarlo a la métrica de árbol más próxima.

El punto correspondiente del Dressiano sería:

$$[-41 : -47 : -34 : -28 : -18 : -39 : -41 : -45 : -47 : -34]$$

Ahora que ya tenemos 5 hojas que cumplan nuestras hipótesis, construimos el árbol con el algoritmo usado en la demostración del Teorema 6.8.

Primero tenemos que escoger 3 hojas (i,j,k) tales que maximicen: $d_{ik} + d_{jk} - d_{ij}$. En nuestro caso $(i,j,k) = (2,3,1)$, obteniendo el árbol $\tau' = \{1,3,4,5\}$ de la Figura 6.6.

La quinta hoja la tenemos que añadir en algún punto de l_4, obteniendo la Figura 6.7.

Con la información que tenemos podemos hallar los valores de a, b y c:

$$
\begin{aligned}
l_4 &= a + c = 29 \\
d_{23} &= b + c = 18
\end{aligned}
$$

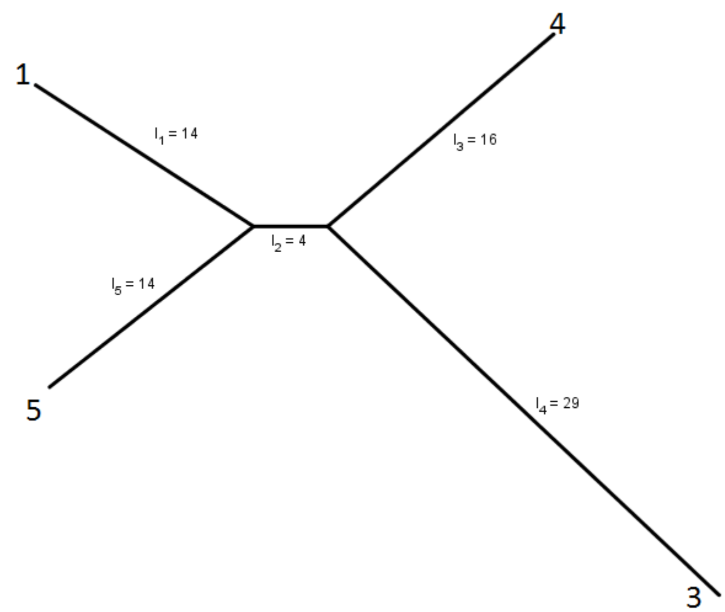


Figura 6.6: Árbol formado por las hojas 1, 3, 4 y 5

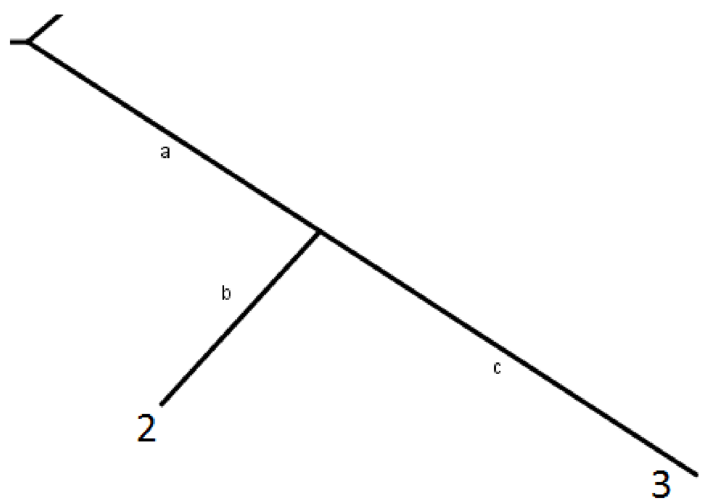


Figura 6.7: Añadimos la quinta hoja en algún punto de l_4

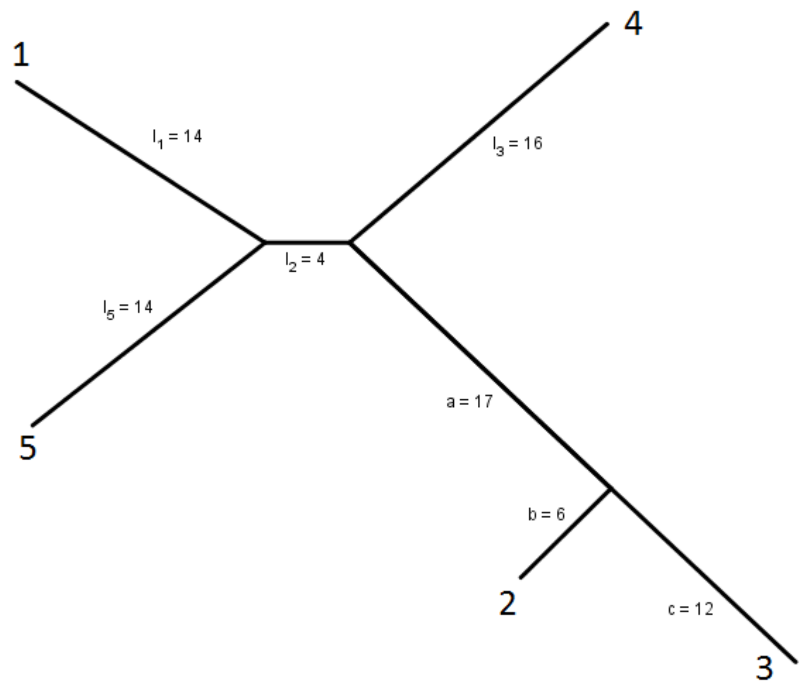


Figura 6.8: Finalmente, terminamos de construir el árbol

$$d_{24} = a + b + 16 = 39$$

Así, obtenemos los valores $a = 19$, $b = 6$, $c = 12$ y obtenemos el árbol de la Figura 6.8.

Bibliografía

[1] ARDILA, FEDERICO; KLIVANS, CAROLINE J. «The Bergman complex of a matroid and phylogenetic trees». J. Combin. Theory Ser. B 96 (2006).

[2] CZUBARA, CRISTIAN CARLOS «Hipersuperficies Tropicales Singulares» Universidad de Buenos Aires, Noviembre 2011.

[3] ESCAMILLA, JUAN F.; CASTILLO MONTES, MAYRA; DUARTE BEZA, SAUL «Interpretacion geométrica de algunas variedades de Schubert» Revista de Matematica: Teoría y Aplicaciones 4(1): 5–19 (1997).

[4] ESQUIVEL, JORGE «¿Que es la taxonomía?» [Online, accedido 18 Feb. 2017]
«http://biologiaesasombrosa.blogspot.com.es/2012/11/que-es-la-taxonomia.html».

[5] MACLAGAN, DIANE; STURMFELS, BERND «Introduction to tropical geometry». Graduate Studies in Mathematics, 161. American Mathematical Society, Providence, RI, 2015.

[6] MIKHALKIN, GRIGORY «Tropical geometry and its applications». International Congress of Mathematicians. Vol. II, 827–852, Eur. Math. Soc., Zürich, 2006.

[7] RICHTER-GEBERT, JÜRGEN; STURMFELS, BERND; THEOBALD, THORSTEN «First steps in tropical geometry». Idempotent mathematics and mathematical physics, 289–317, Contemp. Math., 377, Amer. Math. Soc., Providence, RI, 2005.

[8] SÁNCHEZ PRIETO, DIANA LORENA «Introducción a la Geometría Tropical» Trabajo de Grado, Bogotá - Colombia, Junio 2010.

[9] SIMON, IMRE «On semigroups of matrices over the tropical semiring». RAIRO Inform. Théor. Appl. 28 (1994), no. 3-4, 277–294.

[10] SPEYER, DAVID E. «Tropical linear spaces». SIAM J. Discrete Math. 22 (2008), no. 4, 1527–1558.

[11] SPEYER, DAVID; STURMFELS, BERND «The tropical Grassmannian». Adv. Geom. 4 (2004), no. 3, 389–411.

[12] TORRES ALCARAZ, CARLOS «De la matemática clásica a la matemática moderna: Hilbert y el esquematismo kantiano». Diánoia vol.54 no.63, México. Noviembre 2009.

[13] YU, JOSEPHINE; YUSTER, DEBBIE «Representing tropical linear spaces by circuits» Formal Power Series and Algebraic Combinatorics, Nankai University, Tianjin, China, 2007.

[14] «mtDB - Human Mitochondrial Genome Database» [Online, accedido 18 Feb. 2017] «http://www.mtdb.igp.uu.se/».