APUNTES DE CONVEXIDAD Y OPTIMIZACIÓN.

Incluye demostraciones propuestas en clases (en azul).

Source: `https://github.com/soyfode/matematicas/tree/master/investmat/src/convexOpt`

Christian Limbert Paredes Aguilera

2023

1

Introducción a la optimización convexa

1.1 Introducción

$$
(P) \begin{cases} \min & f_0(x) \\ \\ s.a. & f_1(x) \leq b_1 \\ & f_2(x) \leq b_2 \\ & \vdots \\ & f_m(x) \leq b_m. \end{cases}
$$

Donde,

$$
\begin{array}{ll} f_i: & \mathbb{R}^n \to \mathbb{R} \\ f_0: & \text{Función objetivo.} \\ f_j: & \text{Función Restricción donde } j = 1, \ldots, m. \end{array}
$$

El objetivo de (P) es encontrar x^* el optimo (arg min) que cumpla:

$$f_0(x^*) \leq f_0(x),\ \forall x \in \mathbb{R}^n / f_j(x) \leq b_j,\ j = 1, \ldots, m.$$

Los Puntos factibles son los $x \in \mathbb{R}^n / f_j(x) \leq b_j,\ j = 1, \ldots, m.$

Cuando el problema sea de la forma convexa se llama optimización convexa. Al final, la habilidad es identificar las restricciones y convertirlas a convexas.

- Las funciones objetivos en economía se les puede llamar funciones de coste.
- Multiplicamos las desigualdades por (-1) para darle la forma que más nos convenga darle al problema de optimización.
- Maximizar será lo mismo que minimizar. En nuestro caso minimizaremos las funciones.
- Al valor $f_0(x^*)$ se le llamará valor optimo.
- En $f_i : \mathbb{R}^n \to \mathbb{R}$ existirán algunas funciones el cual su dominio sera "tramposo".

Nota 1.1. *Imaginemos que tenemos*

$$\{\min \quad f(x)$$

$$\{\min \quad f_0^2(x)$$

- *Si las función f_0 es positiva las dos formas son equivalentes.*
- *El valor optimo no será el mismo, pero el punto optimo lo será, ya que las funciones son monótomas crecientes.*
- *Si el valor al cuadrado simplificará la solución, entonces podemos utilizarla. Esto nos permite que si no tengamos una función convexa podamos convexificarla.*

Notación 1.1 Podemos escribir Ax como

$$\begin{aligned} Ax &= \underbrace{\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix}}_{A^1} x_1 + \underbrace{\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix}}_{A^2} x_2 + \cdots + \underbrace{\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{kn} \end{pmatrix}}_{A^n} x_n \\ &= x_1 A^1 + x_2 A^2 + \cdots + x_n A^n. \end{aligned}$$

A^1 = A super 1 como columna, y A_1 = A super 1 como fila.

Ahora, en términos de filas. Si escribimos los vectores A en columna

$$A = \begin{pmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_k^T \end{pmatrix}$$

Donde,

$$Ax = \begin{pmatrix} A_1^T x \\ A_2^T x \\ \vdots \\ A_n^T x \end{pmatrix} = \begin{pmatrix} \langle A_1^T, x \rangle \\ \langle A_2^T, x \rangle \\ \vdots \\ \langle A_n^T, x \rangle \end{pmatrix}$$

Ejemplo 1.1 Sean $A \in \mathcal{M}_{k\times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathcal{R}^k$.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \qquad \mathbf{x}^T = (x_1, x_2, \ldots, x_n).$$

El problema será una minimización global dada por: $\left\{ \begin{array}{ll} \min : & \|A\mathbf{x} - b\|_2^2 \\ s.a. & \emptyset. \end{array} \right.$

Solución.- Por diferenciabilidad se tiene,

$$f_0(x) = \|Ax - b\|_2^2 = \langle Ax - b, Ax - b \rangle.$$

Intentaremos demostrar el punto donde las parciales de $f_0 = 0$. Para

- Diremos que el un vector cualquiera sera vector columna.
- El subindice $_2$ significa la normal Euclidea. Que es la distancia normal que existe en $\mathbb{R}^2$.

ello, encontraremos

$$\begin{aligned} D_i f_0 &= D_i \left(\langle Ax - b, Ax - b \rangle \right) \\ &= \langle D_i \left(Ax - b \right), Ax - b \rangle + \langle Ax - b, D_i \left(Ax - b \right) \rangle \\ &= 2 \langle Ax - b, D_i \left(Ax - b \right) \rangle. \end{aligned}$$

Veamos la parcial de $D_i \left(Ax - b \right)$.

$$D_i \left(Ax - b \right) = D_i \left(x_1 A^1 + x_2 A^2 + \cdots + x_n A^n - b \right) = A^i.$$

Dado que b que es constante vale cero, y donde todos los suman que no estén las x_i también valen cero. Por lo tanto,

$$D_i f_0 = 2 \langle Ax - b, A^i \rangle.$$

Luego,

$$2 \langle Ax - b, A^i \rangle = 0 \ \forall i = 1, \ldots, n \quad \Rightarrow \quad \langle Ax - b, A^i \rangle = 0, \quad \forall i = 1, \ldots, n.$$

Observemos que,

- En funciones convexas el extremo local será el mínimo global.

$$\overrightarrow{0} = \begin{pmatrix} \langle Ax - b, A^1 \rangle \\ \langle Ax - b, A^2 \rangle \\ \vdots \\ \langle Ax - b, A^n \rangle \end{pmatrix} = \begin{pmatrix} (A_1)^T \\ (A_2)^T \\ \vdots \\ (A_n)^T \end{pmatrix} (Ax - b) = A^T (Ax - b).$$

$$A^T (Ax - b) = \overrightarrow{0} \quad \Leftrightarrow \quad A^T A x = A^T b.$$

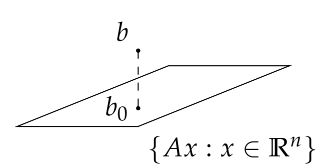
El cual es una ecuación normal.

Por último, veamos los argumentos geométricos. Notemos que,

$$\min \|Ax - b\|_2^2 = d(b_i, Ax)^2$$

Donde Ax tendrá la forma geométrica de un subespacio vectorial (en el caso de $\mathbb{R}^3$ será un plano). Entonces,

- Si $b \in \{Ax : x \in \mathbb{R}^n\} \ \Leftrightarrow \ x^* \in \mathbb{R}^n : Ax^* = b$. El valor optimo es $f_0(x^*) = 0$.
- Si $b \notin \{Ax : x \in \mathbb{R}^n\}$, $f_0(x^*) = d(b, b_0)^2$.



Ahora, cual es el optimo?; es decir cual es el x^*. Para ello,

$$x^* \in \mathbb{R}^n : Ax^* = b_0.$$

Aquí, b_0 está en el plano, si estamos en $\mathbb{R}^3$. ¿Cómo llegamos algebraicamente?:

$$\begin{aligned}
b - b_0 \perp \{Ax : x \in \mathbb{R}^n\} &\Leftrightarrow b - b_o \perp A^i,\ i = 1, \dots, n \\
&\Leftrightarrow \langle b - b_0, A^i \rangle = 0,\ i = 1, \dots, n \\
&\Leftrightarrow \langle b - Ax^*, A^i \rangle = 0,\ i = 1, \dots, n \\
&\Leftrightarrow \langle Ax^* - b, A^i \rangle = 0,\ i = 1, \dots, n \\
&\Leftrightarrow A^T A x^* = A^T b.
\end{aligned}$$

(Las ecuaciones normales vienen dadas por la perpendicularidad.)

2

Conjuntos convexos

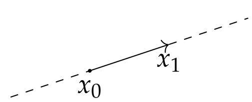
2.1 Conjuntos convexos de $\mathbb{R}^n$

El dominio serán conjuntos convexos o dominio efectivo.

- Cuando λ vale 1 será x_1 cuando valga cero x_0.
- Cuando es positivo irá a la derecha, cuando es negativo hacia la izquierda.
- Toda la recta nos da un concepto que denominamos Afín. Es cualquier punto que este entre x_0 y x_1 del gráfico de arriba.

Definición 2.1 **Lineal.**

$$L(x_0, x_1) := \{x_0 + \lambda(x_1 - x_0) : \lambda \in \mathbb{R}\}$$



Para la convexidad no es necesario tener la linea, solo necesitaremos un segmento definido por:

$$[x_0, x_1] := \{x_0 + \lambda(x_1 - x_0) : \lambda \in [0, 1]\} = \{(1 - \lambda)x_0 + \lambda x_1 : \lambda \in [0, 1]\}$$

- Manejar el concepto de afín con lineas es incomodo, por lo que se utiliza el concepto de combinación afín.
- La diferencia entre espacio vectorial y espacio afín es que el espacio afín esta desplazado; es decir, no necesariamente pasa por el cero como en un subespacio vectorial.

Definición 2.2 Sea $A \subseteq \mathbb{R}^n$. Se dice **Afín**, si $\forall x, y \in A$ se tiene que la $L(x, y) \subseteq A$. (Subespacios vectoriales desplazados).

- Un circulo no es afín ya que la linea es infinita.
- Un plano podría ser Afín.

- La recta es afín.
- Todo $\mathbb{R}^n$ es afín.
- Un único punto también es afín, dado que $x = y$.

- $(1-\lambda)x_0 + \lambda x_1$ es una combinación lineal de x_0 y x_1. Donde $(1-\lambda)+\lambda = 1$.
- Lo demás puntos fuera del segmento son las combinaciones lineales de x_0 y x_1.

Definición 2.3 Una **combinación afín** de los vectores $\{x_1, x_2, \ldots, x_k\}$ es un vector de la forma

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k.$$

tal que

$$\sum_{i=1}^{k} \lambda_i = 1.$$

- Este conjunto es estable para combinaciones lineales, muy similar al concepto de subespacio vectorial.

Teorema 2.1 A es afín sii A contiene toda combinación afín de sus puntos.

Demostración.- Primero, tomemos puntos arbitrarios $\{x_1, x_2, \ldots, x_k\}$ en A tal que

$$z = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k$$

donde $\sum\limits_{i=1}^{k} \lambda_i = 1$.

Ahora, consideremos dos puntos x_i, x_j de z. Dado que A es afín, entonces $L(x_i, x_j) \subseteq A$, para todo x_i, x_j. Esto implica que z está en A. Intuitivamente, si

$$\lambda_1 x_1 + \lambda_2 x_2, \quad \lambda_3 \lambda_3 + \lambda_4 \lambda_4, \quad \ldots, \quad \lambda_{k-1} x_{k-1} + \lambda_k x_k. \qquad \text{con } \sum_{i01}^{k} y_i = 1.$$

están en A. Entonces, z tendrá que estar en A.

Para demostrar la otra implicación, tomemos dos puntos cualesquiera x_1 e x_k en A. Queremos demostrar que

$$L(x_1, x_k) = \{x_1 + \lambda(x_k - x_1) : \lambda \in \mathbb{R}\}$$

está contenido en A. Por el hecho de que x_1 e x_k están en A, podemos considerar la línea $L(x_1, x_k)$. Cualquier punto en esta línea se puede expresar como

$$z = x_1 + \lambda(x_k - x_1),$$

donde $\lambda \in \mathbb{R}$. Ahora, notemos que $\lambda + (1-\lambda) = 1$, lo cual es la condición de combinación afín. Y dado que A contiene toda combinación afín de sus puntos, esto implica que z está en A. Por lo tanto, A es afín. ■

Nota 2.2. *La definición de subespacio se refiere a tomar dos escalares y dos vectores, realizar la* *combinación lineal, donde esta combinación lineal no se saldrá del conjunto dado.*

Notación 2.1 La suma de Minkowski es la operación de conjuntos; es decir, si $A, E \subseteq \mathbb{R}^n$. Entonces,

$$A = x_0 + E = \{x_0 + e : e \in E\} \quad \text{o} \quad E = A - x_0 = \{a - x_0 : a \in A\}.$$

Es sencillamente trasladar los puntos del plano y desplazarlos o moverlos.

Teorema 2.2 $A \subseteq \mathbb{R}^n$ es afín sii existe un subespacio vectorial $E \subseteq \mathbb{R}^n$ tal que $A = x_0 + E$ para todo $x_0 \in A$.

Demostración.- Supongamos que A es afín y fijamos $x_0 \in A$. Intentaremos probar que $E = A - x_0$ es un subespacio de $\mathbb{R}^n$, esto es equivalente a decir que:

$$\lambda, \mu \in \mathbb{R}, e_i, e_2 \subseteq E \quad \Rightarrow \quad \lambda e_i + \mu e_2 \in E.$$

Probemos que $\lambda e_1 + \mu e_2 \in E$; en otras palabras, probaremos que $\lambda e_1 + \mu e_2$ es $a - x_0$.

$$\begin{aligned} \lambda e_1 + \mu e_2 &= \lambda(a_1 - x_0) + \mu(a_2 - x_0) \\ &= \lambda a_1 + \lambda a_2 - \lambda x_0 - \mu x_0 \\ &= \lambda a_1 + \lambda a_2 - \lambda x_0 - \mu x_0 + x_0 - x_0 \\ &= \lambda a_1 + \lambda a_2 + (1 - \lambda - \mu) x_0 - x_0. \end{aligned}$$

Observemos que $\lambda a_1 + \lambda a_2 + (1 - \lambda - \mu)x_0$ está en A, dado a que $\lambda + \mu + (1 - \lambda - \mu) = 1$. Por lo tanto,

$$A - x_0 = E.$$

Es un subespacio vectorial.

Ahora, para demostrar que A es afín, probaré que cualquier combinación afín de elementos de A sigue estando en A (Teorema 1.1). Sean,

$$\{a_1, a_2, \ldots, a_k\}, \ \lambda_1, \lambda_2, \cdots, \lambda_k : \sum \lambda_i = 1.$$

De donde,

$$\begin{aligned} \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k &= \lambda_1(e_1 - x_0) + \lambda_2(e_2 - x_0) + \cdots + \lambda_k(e_k - x_0) \\ &= \lambda_1 e_1 + \cdots + \lambda_k e_k + \left(\sum_{i=1}^{k} \lambda_i\right) x_0 \end{aligned}$$

Observemos que $\lambda_1 e_1 + \cdots + \lambda_k e_k$ es una combinación lineal afín el cual existe en E y por definición, $\left(\sum_{i=1}^{k} \lambda_i\right) = 1$. Por lo tanto,

$$E + x_0 = A.$$

■

Definición 2.4 **Envoltura Afín.**

La envolura afín de B, $\mathrm{Aff}(B)$, es el menor conjunto afín que contiene a B. Esto implica que es el conjunto de las combinaciones afines de elementos de B o es la intersección de los conjuntos afines que contienen a B

Definición 2.5 Si A es Afín se llama "dimensión afín de A" a la dimensión de su espacio vectorial.

- Dimensión 0 un punto.
- Dimensión 1 una recta.
- Dimensión 2 una plano.

Ejemplo 2.1 Dado $C \in \mathbb{R}^n$ afín. Siempre existirán una matriz $A \in \mathcal{M}_{p\times n}$ y $b \in \mathbb{R}^p$ tal que

$$C = \{x \in \mathbb{R}^n : Ax = b\}.$$

Solución.- El conjunto lineal asociado será el núcleo de la aplicación lineal. Es decir,

$$E = \{e \in \mathbb{R}^n : Ae = 0\},$$

cualquier solución de $x_0 \in C$ de modo que $Ax_0 = b$. Tomando un punto de C y otro de E, tenemos

$$A(x_0 + e) = Ax_0 + Ae = b + 0 = b.$$

Por lo tanto,

$$C = \{x \in \mathbb{R}^n : Ax = b\} = x_0 + E.$$

Así, el conjunto afín no es más que el traslado del espacio vectorial.

Definición 2.6 **Topología de $\mathbb{R}^n$.**

Sean $A \subseteq \mathbb{R}^n$ y $a \in \mathbb{R}^n$:

1) $a \in A$ está en el interior de A $\left(a \in \mathrm{int}(A) \text{ o } a \in \mathring{A}\right)$, cuando existe $\delta > 0$ tal que $B(a,\delta) \subseteq A$.

$$B(a,\delta) = \{x \in \mathbb{R}^n : \|x-a\|_2 \le \delta.\}.$$

2) A se dice abierto si $A = \mathrm{int}(A) = \mathring{A}$.

3) Decimos que $c \in \mathbb{R}^n$ está en el cierre (o clausura) de A, cuando $\exists \{a_n\} \in A | a_n \to c$.

4) Decimos que A es cerrado cuando $A = \overline{A}$ donde

$$\{x \in \mathbb{R}^n : x \text{ está en el cierre de A}\}.$$

- El concepto de punto interior es importante, ya que podemos acercarnos al punto a de todas las direcciones.
- Si es un punto relativo interior nos acercaremos por todos los lados del conjunto.
- El punto de adherencia o clausura es un punto el cual me puedo acercar de alguna forma.

1) En $\mathbb{R}^2$ será un circulo y en $\mathbb{R}^3$ será una esfera.

2) Son las bolas que están completamente dentro del conjunto. Es decir, no tienen puntos frontera.

3) Es cualquier bola de c que corta al conjunto o los puntos que contienen a toda su frontera.

4) El cierre son los puntos interior y los puntos frontera.

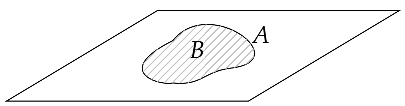
5) Cualquier bola está en el interior cómo en el exterior del conjunto.

5) Se llama frontera de A, ∂A a la intersección $\overline{A} \cap \left(\overline{\mathbb{R}\backslash A}\right) = \overline{A} \backslash \operatorname{int}(A)$ (Cualquier bola estará una parte en el interior y otra en el exterior del conjunto).

6) $a \in \operatorname{relint}(A)$ si existe $\delta > 0$ tal que $B(a,\delta) \cap Aff(A) \subseteq A$.

6) Imaginamos un corte transversal para proyectar una imagen.

Ejemplo 2.2 Dibujemos un plano ($\mathbb{R}^3$)



- A es la frontera.
- El objetivo será encontrar el punto optimo de un esfera que está proyectada en este plano.
- B es el interior con la frontera.
- El conjunto tendrá que ser convexo.

Veamos algunas propiedades de este conjunto.

1) A es cerrado | Cualquier punto que ponga en B me puedo acercar por puntos de B.

2) B cerrado.

3) $\mathring{A} = \emptyset$ | Si yo ponga una bola $\mathbb{R}^3$, se saldrá del conjunto A.

4) $\mathring{B} = \emptyset$ | Ya que no existirá en el plano ninguna esfera.

5) $\operatorname{relint}(A) = \emptyset$; $\operatorname{relint}(B) = B\backslash A$.

- La única diferencia entre combinación convexa y afín es que la combinación convexa es positiva.

Definición 2.7 **Combinación convexa.**

Sean $x_1, x_2, \ldots, x_k \in \mathbb{R}^n$ y $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ tales que

$$\lambda_i \geq 0 \text{ y } \sum_{i=1}^{k} \lambda_i = 1.$$

Al vector

$$\sum_{i=1}^{k} \lambda_i x_i = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k$$

se le llama combinación convexa de los puntos $\{x_1, \ldots, x_k\}$.

- En particular, las combinaciones convexas son los segmentos.

Observación 2.1 La definición para 2 puntos $\{x_1, x_2\}$ nos da las combinaciones convexas,

$$\lambda x_1 + (1-\lambda)x_2, \qquad \lambda \geq 0, (1-\lambda) \geq 0 \iff \lambda \in [0,1].$$

Esto es el segmento,

$$\{\lambda x_1 + (1-\lambda)x_2 : \lambda \in [0,1]\} = [x_1, x_2].$$

Nos quedamos con el segmento que los une, eso nos permitirá utilizar las propiedades de los números reales. Por lo que podremos realizar análisis.

- Que C sea cerrado para las combinaciones convexas, quiere decir que no me salgo del conjunto.

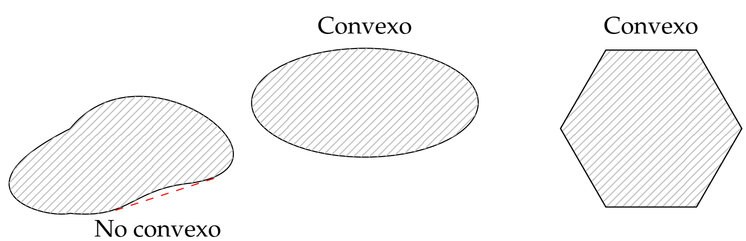
Definición 2.8 **Convexo.**

Un conjunto $C \in \mathbb{R}^n$ se dice convexo cuando C contiene las combinaciones convexas de sus puntos, si y sólo si

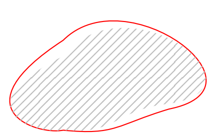
$$\forall x_1, x_2 \in C \Rightarrow [x_1, x_2] \subseteq C.$$

Un conjunto es convexo si dados dos puntos el segmento que los une se queda adentro.

Ejemplo 2.3



Del gráfico 1) ¿Cuál es el menor conjunto convexo que lo contiene?



Definición 2.9 Se llama **envoltura convexa** de A al menor conjunto convexo que lo contiene o a la intersección de todos los convexos que contienen a A, denotado por $\operatorname{co}(A)$. También es equivalente a decir que

$$\operatorname{co}(A) = \{\text{Combinación convexa de puntos de A.}\}$$

Ejercicio 2.1 Demostrar que la intersección de conjuntos convexos es convexo.

Demostración.- Demostremos por contradicción. Sean C_1 y C_2 dos conjuntos convexos. Y sea

$$C = C_1 \cap C_2.$$

no convexo. Esto significa que existen x e y tales que

$$\{\lambda x + (1-\lambda)y : \lambda \in \mathbb{R}\} \not\subseteq C.$$

Supongamos ahora que x e y están en C. Cómo ambos C_1 y C_2 son convexos, el segmento definido debe estar en ambos conjuntos. Es decir,

$$\{\lambda x + (1-\lambda)y : \lambda \in \mathbb{R}\} \subseteq C.$$

Lo que contradice nuestra suposición inicial. Por lo tanto, C es convexo. ■

> El concepto de cono será importante, de hecho existe un tipo de optimización basado en conos convexos. Uniremos la propiedad de ser cono y convexos, pero aunque no sean convexos son importantes.

Definición 2.10 **Cono.**

Un conjunto $C \subseteq \mathbb{R}^n$ se llama cono si y sólo si

$$\lambda x \in C \text{ si } x \in C, \ \lambda \geq 0.$$

Contiene los rayos que pasan por el cero e intersecan a un punto dado.

Propiedad 2.1 **Propiedades de los conos.**

a) Un cono siempre contiene al origen.

b) La envoltura cónica de un conjunto es $\text{con}(A) = \{\lambda : \lambda \geq 0, a \in A\}$. La intersección de todos los conos que contiene a A.

c) Un cono C es convexo si y sólo si

$$\lambda_1, \lambda_2 \in C \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in C, \ \forall \lambda_1, \lambda_2 \geq 0.$$

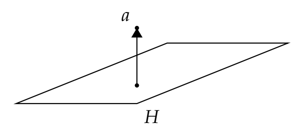
d) C es un cono convexo si

$$\sum_{i=1}^{m} \lambda_i x_i \in C$$

para $\lambda_i \geq 0$.

Cuando uno estudia convexos la principal herramienta es el hiperplano. Es un caso particular de convexo.

- Hiperplano:



- En $\mathbb{R}$ los hiperplanos son rectas.
- En $\mathbb{R}^n$, los hiperplanos serán uno menos de dimensión.

Definición 2.11 **Hiperplano.**

$H \subseteq \mathbb{R}^n$ es un **hiperplano** si existe $a \in \mathbb{R}^n \setminus \{0\}$ tal que

$$H = \left\{x \in \mathbb{R}^n : \langle a, x\rangle = a^T x = 0\right\} = a^{\perp}.$$

Proposición 2.1 $H \subseteq \mathbb{R}^n$ es un hiperplano si y sólo si, H es un subespacio de dimensión $n-1$.

Demostración.- Primero, supongamos que H es un hiperplano. Entonces, por definición, existe un vector $a \in \mathbb{R}^n \setminus \{0\}$ tal que

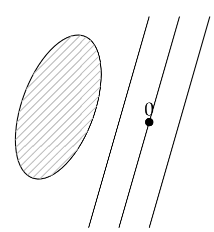
$$H = \left\{x \in \mathbb{R}^n : a^T x = 0\right\} = a^{\perp}$$

Esto significa que H es el conjunto de todos los vectores que son ortogonales a a. Recordemos que cualquier múltiplo escalar del primer vector también es ortogonal al segundo; además, si dos vectores son ortogonales a otro, entonces la suma de los dos primeros vectores también será ortogonal al tercer vector. Es decir, la ortogonalidad preserva las operaciones de suma y multiplicación por escalares. Por lo tanto, H es un subespacio de $\mathbb{R}^n$. Ahora bien, como a no es el vector cero, el conjunto $\{a\}$ es linealmente independiente (ya que no hay otros vectores), y por lo tanto forma una base para un subespacio de $\mathbb{R}^n$. Como este subespacio es ortogonal a H, la dimensión de H debe ser $n-1$.

Ahora supongamos que H es un subespacio de $\mathbb{R}^n$ de dimensión $n-1$. Entonces, existe un subespacio de $\mathbb{R}^n$ que es ortogonal a H y tiene dimensión 1. Este subespacio tiene una base formada por un único vector, digamos a. Entonces, para cualquier vector $x \in H$, tenemos que $a^T x = 0$, lo que significa que x es ortogonal a a. Por lo tanto, podemos escribir $H = \{x \in \mathbb{R}^n : a^T x = 0\}$, lo que significa que H es un hiperplano. ■

Observemos que la recta que pasa por el cero estará definida por:

$$\left\{x \in \mathbb{R}^n : \langle a, x\rangle = a^T x = 0\right\} = a^{\perp}.$$



Y todos los hiperplanos que serán paralelos a esa recta estarán dadas por:

$$\left\{x \in \mathbb{R}^n : \langle a, x\rangle = a^T x = b\right\}.$$

Esto, nos dará dos semiespacios dados por:

$$\left\{x \in \mathbb{R}^n : a^T x \leq b\right\}$$

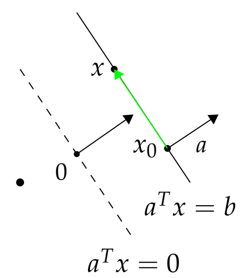
- Un hiperplano en $\mathbb{R}^2$ será sencillamente las rectas.

$$\left\{x \in \mathbb{R}^n : a^T x > b\right\}$$

Esto nos divide el espacio en dos trozos. Que es la estrategia fundamental de análisis de datos. Por ejemplo cuando marcamos con lineas cuando existen datos por arriba y por abajo.

- Esa recta me va a definir dos semiespacios uno al lado del otro.
- Nos interesará desplazar esa recta que contiene al 0.
- La b nos dará una notación de distancia entre los hiperplanos.

Ejemplo 2.4



Para decidir en que dirección estará el punto, debemos tomar en cuenta a que lado apunto a, que nos marcará un punto perpendicular a ese conjunto. De donde,

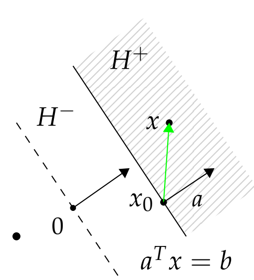
$$\langle (x - x_0), a \rangle = 0.$$

Si queremos en producto matricial se tiene,

$$\begin{aligned} a^T (x - x_0) &= 0 \\ a^T x - a^T x_0 &= 0 \\ a^T x &= a^T x_0 \\ a^T x &= b \end{aligned}$$

- Estas aplicaciones la llaman también aplicaciones del dual.

Ejemplo 2.5



El angulo de $(x - x_0, a)$ esta entre $-90°$ y $90°$. En términos de cosenos sería:

$$\cos\left[\text{ang}(x - x_0, a)\right] \in [0, 1]$$

Luego,

$$0 \leq \cos\left[\text{ang}(x - x_0, a)\right] = \frac{\langle (x - x_0), a\rangle}{\|x - x_0\|_2 \|a\|_2}$$

Por lo tanto,

$$\begin{aligned}
x \in H^+ &\Leftrightarrow \langle (x - x_0), a\rangle = \cos\left[\text{ang}(x - x_0), a\right] \|x - x_0\|_2 \|a\|_2 \geq 0 \\
&\Leftrightarrow \langle (x - x_0), a\rangle \geq 0 \\
&\Leftrightarrow \langle x, a\rangle \geq \langle x_0, a\rangle \\
&\Leftrightarrow a^T x \geq a^T x_0 \\
&\Leftrightarrow a^T x \geq b
\end{aligned}$$

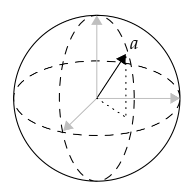
Ahora, si $x \in H^-$, entonces $a^T x \leq b$.

> Esto es el canon de una aplicación lineal. Para poder separar cosas con hiperplanos, lo único que tengo que hacer es aplicar una aplicación lineal y una desigualdad.

2.2 Bolas Euclideas

Tenemos que,

$$\begin{aligned}
B(c, r) &= \{x \in \mathbb{R}^n : \|x - c\|_2 < r\} \\
&= \left\{x \in \mathbb{R}^n : (x - c)^T (x - c) < r^2\right\}.
\end{aligned}$$



Ejercicio 2.2 Demostrar que $B(c, r)$ es convexo.

Demostración.- Sean x_0, x_1 en $B(c, r)$ y $\lambda \in [0, 1]$. Demostraremos que

$$(1 - t)x_0 + \lambda x_1$$

esta también en $B(c, r)$. Primero, notemos que

$$\|x_0 - c\|_2 < r \quad \text{y} \quad \|x_1 - c\|_2 < r.$$

Luego, por la definición de convexidad, y por la desigualdad triangular,

$$\begin{aligned}
\|\left[(1 - \lambda)x_0 + \lambda x_1\right] - c\|_2 &= \|\lambda(x_0 - c) + (1 - \lambda)(x_1 - c)\|_2 \\
&\leq \lambda\|x_0 - c\|_2 + (1 - \lambda)\|x_1 - c\|_2 \\
&< \lambda r + (1 - \lambda) r.
\end{aligned}$$

Por lo tanto,

$$\|(1-\lambda)x_0 + \lambda x_1 - c\|_2 < r.$$

Concluimos que, $B(r,c)$ es convexo. (La demostración se basó en el libro de Boyd). ■

Propiedad 2.2 Mediante la suma de Minskowsky tenemos,

$$B(c,r) = c + rB(0,r).$$

- Si la bola está al rededor del cero u otro punto, la bola es la misma. Esto en distancias no tiene porque ser cierto.
- Todas las bolas que podamos dibujar podremos representarlos con centro cero, ya sean grandes o pequeñas.

2.3 Elipsoides

$$\begin{aligned}
\mathcal{E} &= \left\{x \in \mathbb{R}^n : (x-c)^T P^{-1}(x-c) \le 1\right\} && (1)\\
&= c + \left\{y \in \mathbb{R}^n : y^r P^{-1} y \le 1\right\} \qquad (y = x - c) && (2)\\
&= c + \left\{y \in \mathbb{R}^n : y^T L D L^T y \le 1\right\} && (3)\\
&= c + \left\{Lz : z^T D z \le 1\right\} \qquad (z = L^T y) && (4)\\
&= c + L\left\{z : z^T D z \le 1\right\}. && (5)
\end{aligned}$$

donde

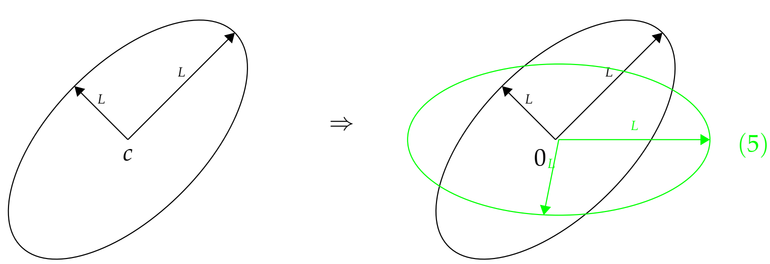
$$P = P^T > 0 \text{ (Simétrica y valores propios } > 0).$$

y

$$P^{-1} = LDL^T, \text{ Si } P \in \mathcal{M}_{n\times n}.$$

donde D es una matriz diagonal con valores propios de $1/P$.

- Cómo es simétrica y tiene valores propios positivos, se puede utilizar la diagonalización y escribir la matriz P cómo producto de una matriz diagonal por dos matrices de cambio que en realidad son ortonormales.
- $z^T Dz$, se puede hacer más grande o mas pequeña.
- Los vectores de L me dan los vectores que apuntan a la elipse.
- $z^T Dz$ es el circulo.
- D serán las curvaturas principales de la elipse.
- La L gira la elipsoide.
- Los elipsiodes se manejan para manejar imagenes donde incluye un objeto.



(3)

(1)

2.4 Bolas generales y conos asociados

Definición 2.12 **Norma.** $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}^+$.

Algunas condiciones:

i) $\|x\| = 0 \Leftrightarrow x = 0$.

ii) Si la norma de un vector se eleva al cuadrado, se esperara que la longitud sea el doble.

$$\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R}.$$

iii) La desigualdad triangular.

$$\|x + y\| \leq \|x\| + \|y\|.$$

Recordemos que,

$$\begin{aligned} B(c,r) &= \{x \in \mathbb{R}^n : \|x - c\| \leq r\} \\ &= \left\{x \in \mathbb{R}^n : d_{\|\cdot\|}(x,c) \leq r\right\}. \end{aligned}$$

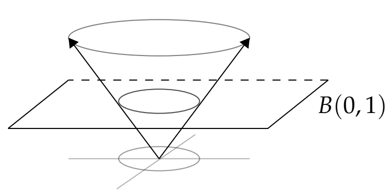
Donde $B(c,r)$ es convexa.

2.4.1 Cono asociado a una norma

- Una vez que conoces la $B(0,1)$ se conoce todas las demás.
- No se puede derivar en el origen, ya que estará en el pico del cono.
- Podemos analizar todo lo que está dentro del cono,
- A esto se le llama epigrafo de la función, es decir dibujar la función y pintar todo lo que hay arriba.
- La definición del cono es cerrada.

Definición 2.13

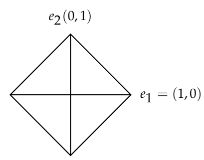
$$C_{\|\cdot\|} = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\| \leq t\}.$$



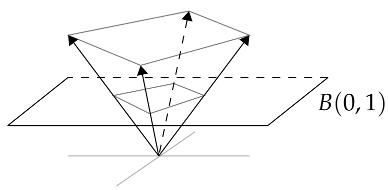
$$\left\{(x,t) : t = 1 \cap C_{\|\cdot\|}\right\} = \overline{B(x,t)}.$$

Tipos de norma

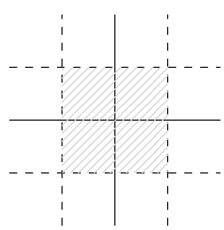
a) $\|\cdot\|_1 : \mathbb{R}^2 \to \mathbb{R}$, $\|(x,y)\|_1 = |x| + |y|$. Si lo definimos en $\mathbb{R}^8$ las sumas serán la suma de sus componentes. Si queremos dibujar $\overline{B(0,1)}$



Donde, e_1 y e_2 se llaman extremales.



b) $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}$, $\|x,y\|_\infty = \max\{|x|,|y|\}$. Compara y coge la más grande. Veamos una vista transversal con respecto del cono:

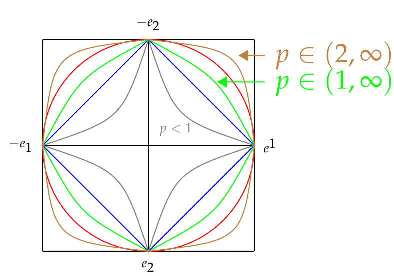


c) $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$, $p \in (1,\infty)$. (Existe una especie de promedio)

$$\|(x,y)\| = \sqrt[p]{|x|^p + |y|^p}$$

En particular:

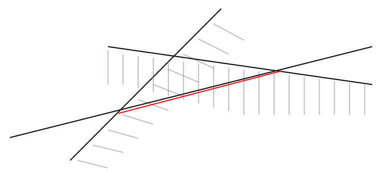
$$\|(x,y)\|_2 = \sqrt{|x|^2 + |y|^2}.$$



- La norma infinito es cómo la madre de todas las demás normas. Representamos con un cuadrado sobre la base canónica.
- La norma 1 sera el rombo que ya dibujamos.
- La Euclídea será un círculo.
- Cuando $p < 1$ no será una función convexa. (No se cumplirá la desigualdad triangular). Ya no será una bola.

2.5 Poliedros

$$P = \left\{ x \in \mathbb{R}^n : a_j^T x \le b_j, j = 1, \dots, n, c_j^T x \le d_j, k = 1, \dots, p \right\}.$$



El poliedro generaliza:

- Linea,
- segmento,
- semiespacio,
- subespacio afín.

Sin embargo no todo conjunto convexo se puede definir cómo un poliedro.

- La idea de los semiespacios afines, de tomar un hiperplano de un lado y el otro, en realidad se puede hacer con hiperplanos o cualquier subespacio afín.
- Un polihedro será prácticamente una cantidad finita de caras. Es decir, son hiperespacio de dimensión $n-1$ que determinan la frontera, tomando desigualdades $a_j^T x \le b_j$. El cual es un semiespacio determinado por la dirección del hiperplano a_j. Y el b_j es una traslación.
- Cuando se pide varias condiciones es la intersección de semiespacios.
- $c_j^T x = d_j$, me fija hiperplanos, que justo corte por un lugar específico.
- Los poliedros son aplicados a optimización lineal.

2.6 Operaciones que conservan la convexidad

Algunas propiedades que conservan la convexidad:

1. Intersección de convexos es convexo.

2. Si tenemos una función $f : \mathbb{R}^n \to \mathbb{R}^m$, f afín. Es decir,

$$(f(x) = Ax + b, A \in \mathcal{M}_{m\times n} \text{ y } b \in \mathbb{R}^m).$$

3. Si tengo un conjunto $C \subseteq \mathbb{R}^n$ convexo. Entonces, $f(c)$ es convexo. Si $B \subseteq \mathbb{R}^n$ convexo, entonces $(f^A(B))$ la anti-imagen de B es también convexo.

Algunos ejemplos particulares de afín:

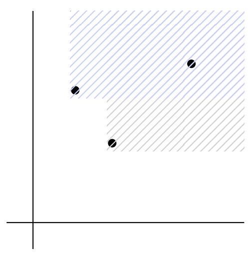
- Las aplicaciones afines cuando trabajamos en $\mathbb{R}^n$ y $\mathbb{R}^m$ son tan sencillas como multiplicar una matriz y sumar un número. Es una aplicación lineal y sumar una constante, es decir trasladar.
- El hecho que funcione para adelante y para atrás, nos permite que para que una función sea convexa yo puedo demostrar que su imagen de f es convexa. Donde se nos simplifica las cosas.
- Si un convexo esta lejos del cero, dado que la traslación es afín, podemos trasladar a cero demostrarlo y llevarlo a su estado original. Así sin perdida de generalidad podemos asumir que el cero está en el conjunto.

- Homotecias: Multiplicar por un escalar.
- Translaciones.
- Proyecciones: Aplicaciones lineales importantes.
- Suma de Minskowsky. Empezamos con dos conjuntos y definimos la suma de Minkowski como el conjunto de los A, B tal que la A está en A y la A en B, lo podemos ver como la imagen aplicación afín de un conjunto convexo.
- Producto de dos conjuntos.

- Si 2., entonces se cumple 3. Lo que se puede ver es que los primeros tres ejemplos son estables, la imagen por una homotecia o la anti-imagen, etc.
- La suma de dos cosas convexas es convexa.
- El producto cartesiano de dos cosas convexas es convexa.
- Estas dos últimas son un poco más difíciles de demostrar porque hay que pensar que funciones afines nos da la suma de Minskowsky y que función afín nos da el producto de dos conjuntos. En realidad, se utiliza la pre-imagen.

2.7 Desigualdad generalizadas

¿Cual de los dos puntos es más importante?. Es los puntos donde están pintados, dependiendo si es máximo o mínimo.



Ahora, cual es mejor de estos dos conjuntos.

Esto se define como:

$$\begin{aligned} x, y \in \mathbb{R}^n, x \leq y \quad &\Leftrightarrow \quad x_i \leq y_i, \quad i = 1, \dots, n \\ &\Leftrightarrow \quad 0 \leq y_i - x_i, \quad i = 1, \dots, n \\ &\Leftrightarrow \quad y - x \in \mathbb{R}^n_+ \\ &\Leftrightarrow \quad y \in x + \mathbb{R}^n_+. \quad \text{(Minkowski)} \end{aligned}$$

Ahora, en vez de utilizar $\mathbb{R}^n_+$ puede utilizarse otro conjunto K. Ahora, ¿Qué propiedades debería tener K, para tener un conjunto con orden?. Necesito asignar una serie de propiedades que tiene $\mathbb{R}^n$.

Por lo que definimos de cono propios

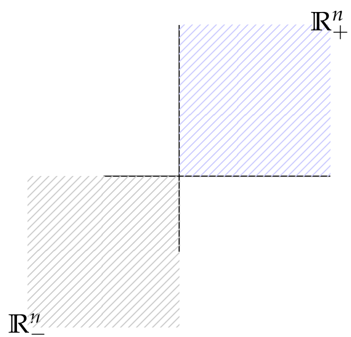
Definición 2.14 $K \subseteq \mathbb{R}^n$ es cono propio si es un cono:

i) Convexo: Cualquier parte de puntos el segmento estará en el mismo conjunto. Buen comportamientos de las SUMAS.

ii) Cerrado: Vamos a definir con el menor o igual.

iii) Sólido ($\text{int}(K) \neq \emptyset$): Que pueda decidir en todo $\mathbb{R}^n$. Para ello, debo tener al menos un punto interior.

iv) Apuntado ($x \in K, -x \in K \Rightarrow x = 0$): No me deja que contenga una recta entera.



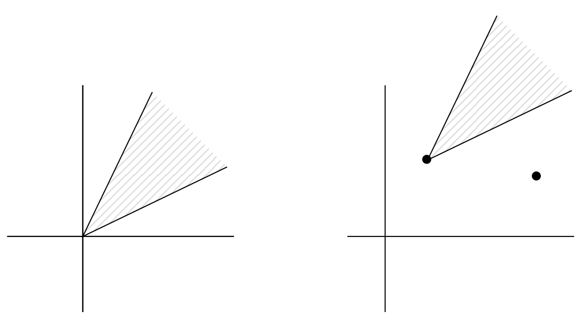
- Está claro que es convexo.
- Ya que no se tiene dos $++$, en $\mathbb{R}^n_+$. Entonces es cerrado.
- Sólido, ya que existe cualquier punto interior. Los bordes no son puntos interiores.
- Apuntado, ya que, cuando tengo un punto tiene opuesto. Básicamente tiene que pasar siempre en cero.

Dado un cono propio $K \subseteq \mathbb{R}^n$ se define $\leq_K$ un orden en $\mathbb{R}^n$ como:

$$x \leq_K y \quad \Leftrightarrow \quad y - x \in K$$
$$\Leftrightarrow \quad y \in x + K.$$

- Cada K que fijemos será una lección de multicriterio.
- Si estoy en $\mathbb{R}^n$ por ejemplo, entonces tengo cinco conos para elegir, y según un criterio de esas 5 variables definiremos cual es mejor o cual es peor.
- Estos criterios lo puedo definir mediante conos.
- Me asegura un sistema sistemático de orden.

Ejemplo: Si tomamos dos puntos. Entonces,



Aquí no puedo asegurar que $x \nleq_K y$.

Ejercicio 2.3 Demuestra que $\leq_K$ es un orden.

i) Reflexiva, $x \leq_K x \ \forall x$.

ii) Transitiva, $x \leq_K y, y \leq_K z \Rightarrow x \leq_K z$.

iii) Antisimétrica, $x \leq_K y, t \leq_k x \Rightarrow x = y$..

iv) Estable para sumas: $x \leq_K y, z \leq_K w \Rightarrow y \pm w$.

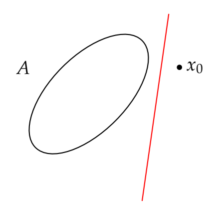
v) Estable para productos positivos: $x \leq_K y \Rightarrow \lambda x \leq_K \lambda y, \lambda > 0$.

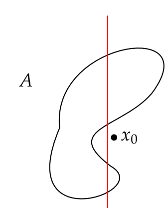
vi) Estable para límites: $x_n \leq_K y, x_n \to x \Rightarrow x \leq_K y$.

Demostración.- Para demostrar debemos utilizar las propiedades de cono propio. Lo que le pido que al orden se comporte bien con las sumas, con los productos por escalares positivos y que se comporte bien con los límites por lo que pido que el cono sea cerrado. ■

2.8 Teorema de separación y extensión

- La idea fundamental es que un conjunto convexo siempre se puede aproximar por tangentes.
- Los hiperplanos nos ayudaran a cortar el espacio, donde a un lado nos dejaran al conjunto y al otro nada.
- Moviendo esos hiperplanos, alrededor del conjunto podemos recuperar toda la forma del conjunto.
- Esto nos dan pie a unos teoremas que nos dice que si tenemos un conjunto y un punto que está fuera, yo siempre puedo encontrar una linea recta que los separa.
- Si el conjunto no es convexo y el punto a separar está en la región "no convexa". Entonces, no podemos separarlas con hiperplanos o semiespacios.





Lema 2.1 **Lema de Zorn.** Si tengo un conjunto $(\mathscr{A}, \leq)$ parcialmente ordenado, tal que si $\mathscr{C} \subseteq \mathscr{A}$ totalmente ordenado $(x, y \in \mathscr{C} \Rightarrow x \leq y$ ó $y \leq x)$, tiene una cota superior en $\mathscr{A}$ $(\exists a \in \mathscr{A} : x \leq a \ \forall x \in \mathscr{C})$. Entonces, $\mathscr{A}$ tiene un maximal $(\exists a \in \mathscr{A} : b \in \mathscr{A},\ b \leq a \Rightarrow b = a)$. $[\mathscr{A} \notin \emptyset]$. ■

- Este lema nos encuentra elementos maximales.
- Maximal significa que no tiene un elemento por encima, si hay otro, entonces es el mismo.
- Sea un conjunto $\mathscr{A}$ parcialmente ordenado, con elementos en fila dentro un conjunto $\mathscr{C}$. Si $\mathscr{A}$ toma la forma de $\mathscr{C}$, siempre estará acotado superiormente.

- Este lema es el mismo enunciado del teorema de abajo, pero no para cualquier subespacio sino para un espacio de co-dimensión. Yo se demostrar que si tengo una función en un hiperplano yo puedo extenderla a una dimensión.
- Con este lema 2.2 podemos realizar operaciones finitas sin necesitad de lema de Zorn.
- Si tenemos un espacio y tenemos un hiperplano y en el hiperplano tenemos las condiciones del teorema. Entonces somos capaces de encontrar una extensión a todo el espacio que sigue siendo lineal que extiende a la g y además está controlada por p.

Lema 2.2 Si un espacio vectorial $H + [x_c]$ y unas funciones $g : H \to \mathbb{R}$ lineal y $p : H + [x_c] \to \mathbb{R}$ lineal. Cumple que $g(h) \leq p(h)\ \forall h \in H$. Entonces, $\exists\ \overline{g} : H + [x_0] \to \mathbb{R}$ lineal, tal que $\overline{g}(h) = g(h)$ y $\overline{g}(x) = g(x)\ \forall x \in H + [x_0]$.

Demostración.- Dado $g : H \to \mathbb{R}$ lineal, tal que $g(h) \leq p(h)\ \forall h \in H$. Quiero definir $\overline{g} : H + [x_0] \to \mathbb{R}$ lineal y que $g(x) \leq p(x)\ \forall x \in H + [x_0]$. Dado que $[x_0]$ es linealmente independiente a $g(h)$, por lo tanto

$$\overline{g}\,(h + \lambda x_0) = \overline{g}(h) + \lambda \overline{g}(x_0) = \alpha = g(h) + \lambda \alpha.$$

Ahora, debemos definir $[x_0]$ para que funcione la cosa. Para ello buscamos $\alpha \in \mathbb{R}$ tal que $g(h) + \lambda x_0 \leq p(h) + \lambda x_0\ \forall h \in H,\ \forall \lambda \in \mathbb{R}$. Esto será equivalente A:

$$\left\{ \begin{array}{lll} g(x) + \alpha & \leq & p(h + x_0),\ \forall h \in H \\ g(h) - \alpha & \leq & p(h - x_0),\ \forall h \in H \end{array} \right. \qquad (1)$$

$$\Updownarrow$$

$$g(x) - p(x - x_0) \leq \alpha \leq p(y - x_0) - g(y)\ \forall x, y \in H.$$

Se puede poner x, y en vez h ya que las desigualdades de (1) son independientes.

Ya que queremos encontrar un número α que cumpla esta última desigualdad. Intentaremos probar que el supremo de lo que tenemos a la izquierda es menor o igual de las cosas que tenemos a derecha. Es decir,

$$\sup \left\{ g(x) - p(x - x_0) \right\}, x \in H \leq \inf \left\{ p(y - x_0) - g(y) \right\}, y \in H.$$

$$\Updownarrow$$

$$g(x) - p(x - x_0) \leq p(y + x_0) - g(y)\ \forall x, y \in H.$$

(Muchas veces para demostrar tomamos algo que no sabemos que es cierta y con equivalencias llegamos a algo que si sabemos que es cierta). Entonces,

$$g(x) + g(y) \leq p(x - x_0) + p(y + x_0)$$

$$\Updownarrow$$

$$g(x + y) \leq p(x - x_0) + p(y + x_0)$$

Dado que $g(x + y)$ está controlada por la p, se tiene

$$g(x + y) \leq p(x + y)$$

Cómo p está definida en todo el espacio y es sublineal, obtenemos

$$\begin{array}{rcl} g(x + y) & \leq & p(x + y) \\ & = & p(x - x_0 + y + x_0) \\ & \leq & p(x - x_0) + p(y + x_0). \end{array}$$

Esta última desigualdad es cierta, por lo que queda demostrada el lema. ■

- Extender funciones lineales es sumamente sencillo: Por ejemplo si estamos en $\mathbb{R}^3$ y tenemos una función lineal definida en el plano, lo único que tenemos que hacer es tomar una base del plano, completarla a la base del espacio y en ente caso como es $\mathbb{R}^3$ y $\mathbb{R}^2$, sería solamente un vector. Entonces, definimos una aplicación lineal que coincida con g en el plano y en el vector que sobra le damos el valor cero.

Teorema 2.3 **Teorema de Hahn-Banach (Extensión.)** Sea E un espacio vectorial y una función $p : E \to \mathbb{R}^+$ sublineal. Sea también $E_0 \subseteq E$, $g : E_0 \to \mathbb{R}$ lineal y $g(x) \leq p(x)\, \forall x \in E_0$. Entonces existe $\overline{g} : E \to \mathbb{R}$ lineal, tal que $\overline{g}(x) = g(x)\, \forall x \in E_0$ ($\overline{g}$ extiende a g) y $\overline{g}(x) \leq p(x)$, $\forall x \in E$.

- La idea es extender combinaciones lineales.
- Es **sublineal** si $p(x+y) \leq p(x) + p(y)$ y $p(\lambda x) = \lambda p(x)$ si $\lambda > 0$.
- La función g lineal es controlada por esa función norma p. Quiere decir que es continua en la métrica g de p.
- La idea es la siguiente: tengo un subespacio, una función lineal definida en el subespacio y que está contralada por p.
- Yo tengo una distancia en todo mi espacio, una aplicación lineal definida abajo y puedo extenderla.

Demostración.- Si yo quiero comenzar por una función g, en un espacio pequeño y quiero extenderlo a un espacio grande, debo comenzar por tomar todas las distribuciones.

Empezamos definiendo de objetos, en este caso extensiones como,

$$f = \left\{ (E_1, g_1) : \begin{array}{l} g_1 : E_1 \to \mathbb{R} \text{ lineal} \\ g_1(x) = g(x)\ \forall x \in E_0 \\ E_0 \subseteq E_1 \\ g_1(x) \leq p(x)\ \forall x \in E_1. \end{array} \right\}$$

Lo primero que justifico es el vacío. Es decir,

$$f \notin \emptyset, \text{ ya que } (E_0, g) \in f.$$

- f es no vacío porque la función inicial g y el subespacio E_0 están ahí.
- La idea de este teorema se basa en que tengo una g que la extiendo por muchas ramas, donde cada extensión las puedo extender de nuevo por más ramas. Pero cuando tomo la cadena totalmente ordenado solo me estaré yendo por una rama. Donde las g van coincidiendo a través de las ramas. Básicamente la f es un árbol y la g es una rama bien ordenada.

Ahora, introducimos un orden en esta familia, de forma que una de estas extensiones sea mejor que otra. Dos los dos elementos definidos,

$$(E_1, g_1) \leq (E_2, g_2).$$

Al definir esto, no decimos que todos los elementos se puedan comparar. Lo que decimos es que si yo tomo estos elementos se cumplirán siempre que se cumpla las siguientes condiciones:

- $E_1 \subseteq E_2$.
- $g_1(x) = g_2(x)\ \forall x \in E_1$.

El conjunto

$$(f, \leq)$$

se llama parcialmente ordenado.

Ahora demostraremos que que este conjunto parcialmente ordenado, toda cadena $\mathscr{C} \subseteq \mathscr{A}$ tiene una cota superior, y por el lema de Zorm tiene un elemento maximal.

Si $\mathscr{C} \subseteq f$ es una cadena totalmente ordenado. Lo que queremos encontrar un elemento no necesariamente en $\mathscr{C}$ pero si en f que sea una cota superior. Para ello,

$$\mathscr{C} = \{(E_i, g_i)\}_{i \in I}$$

Ahora, definimos

$$E_{\mathscr{C}} = \bigcup_{i \in I} E_i.$$

Dados dos que yo fije, uno mayor que otro, siempre tendré inclusión. Por eso nos sirve la unión. Porque están contenido uno debajo del otro.

Definimos también

$$g_{\mathscr{C}} : E_{\mathscr{C}} \to \mathbb{R}.$$

Lo que tendrá que coincidir con las g_i. Por lo que,

$$g_{\mathscr{C}}(x) = g_i(x) \text{ si } x \in E_i.$$

Dado que g_i cumplen $g_1(x) \leq p(x)\ \forall x \in E_1$. Entonces, en lo particular

$$g_{\mathscr{C}}(x) \leq p(x)\ \forall x \in E_{\mathscr{C}}.$$

Así,

$$\left(E_{\mathscr{C}, g_{\mathscr{C}}}\right) \in f.$$

Y además,

$$(E_{\mathscr{C}}, g_{\mathscr{C}}) \geq (E_i, g_i)\ \ \forall i \in I.$$

Donde $\left(E_{\mathscr{C}, g_{\mathscr{C}}}\right)$ es una cota superior de $\mathscr{C}$.

¿Qué haremos con ese elemento maximal?. Lo ideal es que $\overline{E}$ sea todo el E, pero podría ser que no. ¿Qué pasa si $\overline{E}$ se encuentra por debajo de E?. En otras palabras, no llegue a extender a todo el espacio vectorial.

Por lo tanto, por el lema de Zorm existe $\left(\overline{E}, \overline{g}\right) \in f$ que es un elemento maximal de f.

Ahora, demostraremos que $\overline{E}$ es efectivamente la E. Para ello, utilizaremos una reducción al absurdo. Supongamos que $\overline{E} \subsetneq E$. Entonces existe un $x_0 \in \overline{E} \backslash E$. Por lo que puedo ampliar $\overline{E}$ de la siguiente forma:

Cuando tomamos $[x_0]$ quiere decir que tomo la linea generada por x_0 (span).

$$\overline{E} \subsetneq \overline{E} + [x_0] \subseteq E.$$

Lo que intentaremos probar por contradicción es que si $\overline{E}$ no fuese E. $\overline{E}$ no puede ser un elemento maximal. Por el lema 2.2, aplicado a $H = \overline{E}$ y a $\overline{g} : E \to \mathbb{R}$. Existe una extensión $\overline{g} : E + [x_0] \to \mathbb{R}$ tal que

$$\overline{g}(x) \leq p(x)\ \forall x \in \mathbb{E} + [x_0].$$

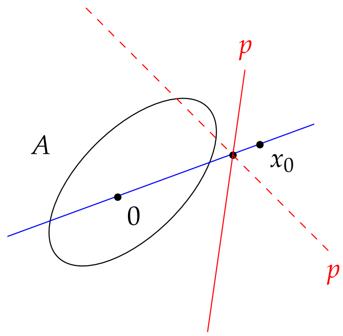
Ahora,

$$\left(\overline{E} + [x_0], \overline{g}\right) \in f \lneq \overline{E}, \overline{g}.$$

Por lo que es imposible pues $\left(\overline{E}, \overline{g}\right)$ es un elemento maximal. Luego, $\overline{E} = E$ y $\left(\overline{E}, \overline{g}\right)$ es la extensión buscada. ■

Ahora el objetivo es establecer los resultados de separación y ver la conexión con el teorema de extensión.

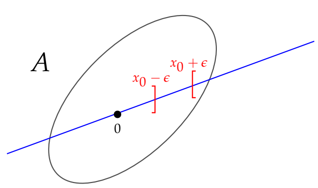
En principio será importante ver un punto interior y separar un punto del conjunto. Tal como lo veremos a continuación:



Nos olvidamos del plano y nos fijamos en la linea. Esta linea es una aplicación lineal. Y por el teorema de Hahn-Banach extiendo. Pero debemos aseguramos que la extensión no se meta en el conjunto, que es precisamente de lo que se encarga p. Esta p a escoger dependerá del conjunto A y la elección de p me permitirá, que cuando yo tome la extensión se extienda fuera de A. Ahora, ¿cuál es la p que se definirá a partir del conjunto A?. Esta p se llama funcional me Minskowsky, y nos hará falta definir el concepto de interior. En dimensión infinita el concepto que mencionaremos a continuación es diferente y en $\mathbb{R}^n$ son lo mismo.

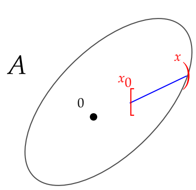
Definición 2.15 Si $S \subseteq E$,

1) $x_0 \in S$ es un punto interior algebraico. Es decir, si $\forall x \in E$, existe un $\epsilon > 0$ tal que $]x_0 - \epsilon, x_0 + \epsilon[\subseteq S$. (Quiere decir que dependiendo de la dirección existirá un ϵ).



2) $S^i = \{x \in S : x \text{ es interior algebraico a } S\}$.

3) x_0 es un punto de clausura algebraica de S, si existe $x \in E$ tal que el intervalo $[x, x_0) \in S$.



4) S^c son los puntos de clausura algebraica de S.

$$S^i = \text{int}(S) \in \mathbb{R}^n$$

$$S^c = \text{ al cierra de } S \text{ en } \mathbb{R}^n \text{ si } S \text{ es convexo.}$$

Lo que necesitamos es definir un punto interior

Definición 2.16 Sea $S \subseteq E$, $0 \in S^i$. Se define el funcional de Minkowski asociado a S cómo:

$$P_S(x) = \inf\{\lambda > 0 : x \in \lambda S\}.$$

- $0 \in S^i$ significa que el cero pertenece al interior algebraico de S.
- La idea de esta definición, es que en vez de tomar la recta, tomaremos una semirecta, que empieza en el cero y se va a más infinito. Cómo el elipsoide es convexo, los bordes serán mi forma de medir (será la p del teorema 2.3), donde hasta el margen valdrá 1, la mitad 1/2 y el doble 2.
- Debemos hacer que se cumple 1) de la definición 2.15. Es decir, que exista un valor ϵ que se quede dentro del conjunto.
- Básicamente lo que nos dice el funcional de Minkowski que está dentro o fuera de un conjunto.

Propiedad 2.3 S convexo, $0 \in S^i$.

1) $S^i = \{x \in E : p(x) < 1\} \subseteq S \subseteq \{x \in E : p_S(x) \leq 1\} = S^c$.

2) $p_S = p_{S^i} = p_{S^c}$

3) $p_S \geq 0$ es sublineal, sii los puntos de p son menores iguales a 1. Y si consigo que sean mayores que uno, las cosas están fuera. En otras palabras, nos permite decidir si estoy fuera o estoy dentro.

4) Reciprocamente, si p es un sublineal y no negativo. Entonces,

$$\{x \in E : p(x) \leq 1\}$$

es convexo y $0 \in S^i$. Además $p = p_S$.

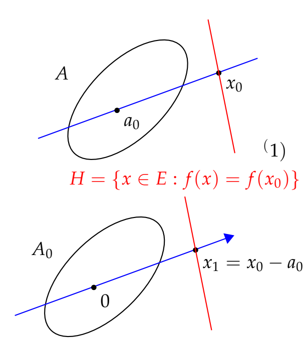
- Los funcionales lineales no negativos están en correspondencia univoca con los conjuntos. Es decir, dame un sublineal encuentro un conjunto y viceversa.)
- Si estamos trabajando con normas, dame una norma y te doy una bola unidad. La bola unidad es el conjunto que define la norma y la norma se define a partir de la bola unidad.

> Esto es el lazo unión entre lo que sería extender funcionales y la separación. Realizaremos ahora cuatro demostraciones utilizando el teorema de Hahn-Banach y el funcional de Minkowski.

Asumimos que el conjunto no tiene borde.

Teorema 2.4 **Teorema de separación.** Sean $A = A^i \neq \emptyset$, $x_0 \in E$, $x_0 \notin A$ y A convexo. Entonces, $\exists f : E \to \mathbb{R}$ lineal, tal que $f(a) \leq f(x_0)$ $\forall a \in A$.

Escrito de otra forma, $A = A^i \neq \emptyset$, convexo, $x_0 \in E \backslash A \Rightarrow \exists f \in E^\#$ con $f(x_0) < f(a)$, $\forall a \in A$.



Demostración.- Tomemos un $a_0 \in A^i$, luego sea un nuevo conjunto $A_0 = A - a_0$. Es decir, tomo el conjunto que tengo convexo y lo traslado.

Ahora como $a_0 \in A^i$, lo que tengo es $0 \in A_0^i$ (1) y $x_1 = x_0 - a_0$.

Definamos el funcional de Minkowski de la siguiente manera:

$$p(x) = p_{A_0}(x)$$

Según el rayo del gráfico, de 0 al borde del conjunto valdrá 1 y por fuera será mayor que 1 (es lo que consigue p para el x_0 fijo). Luego, definamos una función en el span o en el rayo del gráfico de abajo:

$$f[x_1] \to \mathbb{R}$$

de forma que $f(\wedge x_1) = \lambda p(x_1)$ ($\wedge$ múltiplo de x_1). El objetivo de definir esto último es que cuando evalúe esa función en el x_1 me salga un número más grande que 1. Pero los puntos por detrás de 0 nos dará problema, así que lo que quiero demostrar es si $f \leq p$. Y así podemos extender.

Ahora, si $\lambda \geq 0$. Entonces,

$$f(\lambda x_1) = \lambda p(x_1).$$

Gracias a que p es sublineal,

$$f(\lambda x_1) \leq p(\lambda x_1).$$

Luego, si $\lambda < 0$. Entonces,

$$f(\lambda x_1) = \lambda f(x_1) = \lambda p(x_1) \leq 0 \leq p(\lambda x_1).$$

Es decir,

$$f(\lambda x_1) \leq p(\lambda x_1).$$

Y la p está definida en todo el espacio E. ■

- Dual algebraico: $E^{\#}$ = aplicaciones $f : E \to \mathbb{R}$ lineales.
- Nos imaginamos un conjunto sin el borde.
- Lo que haremos es buscar una función f.
- Si tomo el hiperplano H que son los puntos del espacio vectorial donde $f(x) = f(x_0)$ e (valor fijo), que pasa justo por el x_0.
- Lo que me asegura la función f es que ningún punto de los que están en A están en el hiperplano, de hecho todos están a un lado del hiperplano.
- Los puntos que están en A valen estrictamente menor.
- La funcional de Minkowski, cuando voy a la izquierda del cero y multiplico por λ negativos puede ser mucho más estrecha o más grande que a la derecha del cero y en el conjunto. Lo gracioso que siempre λ será positivo.

Teorema 2.5 Aplicando el teorema de Hahn-Banach a $f[x_1] \to \mathbb{R}$ obtenemos $f : E \to \mathbb{R}$ lineal con $f(x) \leq p(x)$, $\forall x \in E$.

Demostración.- Tomemos un punto en $a \in A$ inicial (A^i) donde

$$a - a_0 \in A_0 = A_0^i = \{x \in E : p(x) < 1\}.$$

En particular, quiere decir,

$$p(a - a_0) < 1.$$

Por tanto,

$$p(x_1) = p(x_0 + a_0) \geq 1.$$

(La igualdad de esta última ecuación podría ser porque x_0 está justo en el borde puede estar en la clausura algebraica).

- La función f vista de en la linea que pasa por el cero y x_1 del gráfico de abajo, lo que hace es darle los valores 0 al borde del conjunto valor 1 y todo lo demás.
- Ahora, Hahn-Banach lo extiende cómo sea, con la garantía que tengo de que esa f que extendí, está controlada por p.
- El $a - a_0$ está en el interior relativo
- Recordemos que el funcional de Minkowski de un conjunto es menor que 1 en el interior algebraico. Interior relativo si estamos en $\mathbb{R}$.
- La p controla lo que está adentro y afuera del conjunto.
- Al final la idea es: Tengo un punto fuera ya sea en A o A_0. Elijo un función que es representada por la linea azul del gráfico, Extiendo con Hahn-Banach con el control de la p, evita que si estoy fuera no vaya adentro y contrariamente.

Ahora, veamos que la f separa, por lo que calculamos

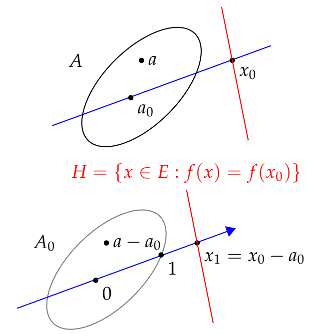
$$f(a-a_0) \leq p(a_0) < 1 \leq p(x_1) = f(x_1) = f(x_0 - a_0).$$

(Realizamos, el calculo en $a - a_0$, porque se hacer la prueba en en el conjunto A_0, luego de conocer el resultado lo podremos trasladar al conjunto A de interés). Por lo tanto, aplicando linealidad

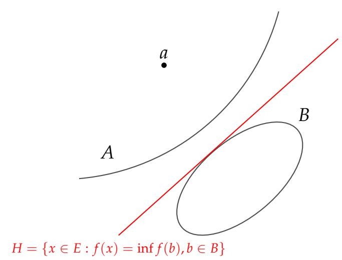
$$\begin{aligned} f(a-a_0) &< f(x_0-a_0) \\ f(a)-f(a_0) &< f(x_0)-f(a_0) \\ f(a) &< f(x_0) \quad a \in A. \end{aligned}$$

■

- En otras palabras, separamos un punto de un conjunto.



Separar un punto de un conjunto parece poca cosa, así que separaremos dos conjuntos.



Teorema 2.6 Sean A y B conjuntos convexos con $A \cap B = \emptyset$ y $A^i \neq \emptyset$. Entonces, existe $f \in E^\#$ lineal, tal que

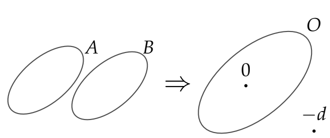
$$f(a^i) < \inf\{f(b),\ b \in B\}$$

para todo $a \in A^i$.

Demostración.- La idea es definir un nuevo conjunto, $O = A^i - B - d$, para cierto $d \in A^i - B$ (observamos que $A^i - B$ es abierto $\Rightarrow$, $o \in \mathring{O}$.) Además, $-d \notin O$. Ya que, si $-d \in O \Rightarrow a^i \in A^i, b \in B$ tal que

$$-d = a^i - b - d \Leftrightarrow a^i = b = 0 \Leftrightarrow a^i = b \Rightarrow A^i \cap B \neq \emptyset.$$

Lo que es una contradicción.



Como A y B son disjuntos. En particular, A^i y B también son disjuntos. Ahora, podemos aplicar el teorema de separación 2.4 al conjunto convexo O y al punto $-d \notin O$, encontrando $f \in E^\#$ tal que

$$f(-d) > f(a^i - b - d), \quad \forall a^i \in A, \forall b \in B.$$

Cuando realicemos apliaciones lineales con sumas o diferencias de Minkowski (unión de conjuntos), y realizar traslaciones, siempre se podrán deshacer por que la función es lineal. Por lo que las sumas y las restas se separan por las propiedades de la función de la linealidad. Por lo tanto,

$$f(a^i) < f(b), \quad \forall a^i \in A^i, \forall b \in B.$$

- El teorema nos dice que tenemos una separación estricta, si encontramos una función f de forma que si lo evalúo en todos los puntos de b y calculo el ínfimo, siempre será menor que la imagen de las f en a^i del interior de A^i.
- Pedimos que al menos uno de los conjuntos tenga un punto interior.
- Si uno de los conjuntos tiene un punto interior, entonces puedo separarlo al igual que separaba un punto de un conjunto.
- El ínfimo, es decir el valor más cercano a B de todos los valores de b, nunca tocará el interior de A.
- Podría coincidir el supremo con el ínfimo. En otras palabras, puede tocar el borde de cada uno pero no al interior.
- Tomamos el hiperplano que cumple exactamente que son todos los puntos que valen $\inf f(b)$.
- Si hacemos H con $\inf f(a)$ saldrá pegado a A.
- Mi objetivo es convertir los dos conjuntos en uno, por lo que los resto.

Podemos tomar supremos e ínfimos de las funciones de la desigualdad, y obtener una desigualdad no estricta. Así nos queda,

$$f(a^i) \leq \inf\{f(b) : b \in B\} = m, \quad \forall a^i \in A^i.$$

Pero el teorema nos dice que existe un desigualdad estricta, por lo cual probaremos que la desigualdad anterior es estricta.

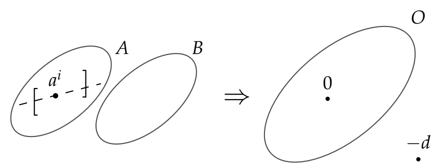
Supongamos que existe un $a^i \in A^i$ tal que

$$f(a^i) = m$$

Como a^i esta en el interior de A^i, podemos construir al rededor de a^i un segmento que estará en el interior de A^i. Es decir, $\exists \lambda > 0$ tal que

$$\left[a^i - \lambda y, a^i + \lambda y\right] \subseteq A^i, \text{ donde } f(y) > 0.$$

que es la definición de interior algebraico (la y podría ser cualquier y interior)..



Luego,

$$\left(a^i + \lambda y\right) \leq m \Leftrightarrow f\left(a^i\right) + \lambda f(y) \leq m.$$

Pero, $f(a^i) + f(y) > m$ dado que $f(a^i) = m$ y $f(y) > 0$.. De donde, es una contradicción. ■

Hasta ahora, demostramos que podemos separar un punto y un conjunto y dos conjuntos, que pedimos que tenga un punto interior. Esto, es importante por que no podremos separar o hacer funciones de Minkowski. Pero tenemos un teorema propio a $\mathbb{R}^n$, que dice lo siguiente:

- En $\mathbb{R}^n$ omitimos que haya puntos interiores. Por lo que la clave de este teorema es como funciona la cosa.
- Cuando r es única, quiere decir que no puedo separar estrictamente los dos conjuntos. De lo contrario se tiene una separación; es decir, un salto entre los dos. Con lo que podemos trabajar.

Teorema 2.7 Sean A y B convexos no vacíos, de forma que ellos no se corta. Es decir, $A \cap B = \emptyset$. Entonces, $\exists r \in \mathbb{R}$, $f \in (\mathbb{R}^n)^{\#}$, tal que $f(a) \leq r \leq f(b)$ $\forall a \in A$ y $b \in B$.

Demostración.- Sea $C \subseteq \mathbb{R}$ convexo y $x_o \notin C$. Análogamente realizaremos la construcción de un punto y conjunto y luego de dos conjuntos. Es decir, puede ocurrir dos cosas:

1. Que la envoltura afín $\mathrm{Aff}(C) = \mathbb{R}^n$. Entonces, C tiene un punto inte-

rior, por lo que por el teorema de separación 2.4, encuentro

$$f \in (\mathbb{R}^n)^\# : f(b) < f(x_0) \forall c \in C^i.$$

Dado que el conjunto C al ser convexo, va a estar siempre entre el interior y el cierre. Entonces, cualquier punto que este en C que no este en C^i es un punto del cierre de C^i. Como el cierre en estos casos,$\mathbb{R}^n$ y convexos, coincide con el cierre algebraico o incluso por continuidad. Como me puedo acercar a todos los puntos del cierre algebraico de $f(c)$ con elementos que están en el interior de C^i y evaluar con la f utilizando la continuidad y la propiedad. Es decir,

$$f(c) \leq f(x_0), \ \forall c \in C.$$

2. Que la envoltura afín $\text{Aff}(C) \subsetneq \mathbb{R}^n$. Sea $c_1 \in C$ y $C - c_i$. Esto último, se trata de tomar la envoltura afín, me sale un subesapcio vectorial que lo traslado al origen. En otras palabras,

$$\text{Aff}(c) - c_1 = E_0 \subsetneq \mathbb{R}^n.$$

Ahora tenemos, un conjunto $C - c_1 \in E_0$, tengo un punto que en principio es $x_0 - c_1 \notin C - c$. De donde puede pasar dos cosas problemáticas: El x_0 puede vivir en E_0 en o que no. Para ello tomemos dos casos:

A) Si $x_0 - c_1 \in E_0$. Entonces, por el teorema de separación 2.4 $\exists f \in (\mathbb{R}^n)^\#$ cumple $f(c - c_1) < f(x_0 - c), \ \forall c \in C^i$. Utilizando linealidad es equivalente a

$$f(c) < f(x_0), \ \forall c \in C^i$$

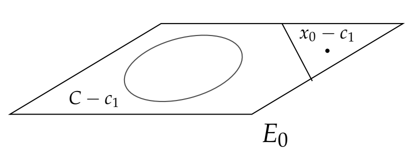
Esto implica que,

$$f(c) \leq f(x_0), \ \forall c \in C.$$

Utilizando el cierre y que las funciones son continuas. Ahora esta f no está definida en todo $\mathbb{R}^n$. En realidad, está definida sólo en el plano, por lo que podemos extenderla f a $\mathbb{R}^n$ y obtenemos $\overline{f} \in (\mathbb{R}^n)^\#$ tal que

$$\overline{f}(c) \leq \overline{f}(x_0), \ \forall c \in C.$$

La extensión me da igual, porque todas las condiciones viven en cero, lo que la función valga fuera de cero, no me interesa. Por la extensión la estoy consiguiente en el espacio de cero.



B) Si $x_0 - c_1 \notin E_0$. No puedo usar el teorema de Hahn-Banach, porque el conjunto que quiero que separar del punto no tiene un punto interior, así que no tengo el funcional de Minkowski para separar. Pero no nos hará falta, ya que es un subespacio vectorial, yo defino una función que en el subesapcio valga cero y en

el punto valga uno; Ya que, los puntos son linealmente independientes.
Tomemos, $f : E_0 + [x_0 - c_1] \to \mathbb{R}$ tal que f es lineal, $f_{|E_0} = 0$ y $f(x_0 + c_1) = 1$. Entonces,

$$f(c - c_1) = 0 < f(x_0 - c_1) = 1 \quad \forall c \in C^i.$$

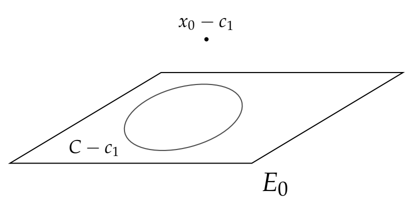
Esto implica que,

$$f(c) < f(x_0), \quad \forall c \in C^i.$$

No hace falta que usemos los puntos interiores, ya que el cierre de c quedará en el hiperplano y ya quedará separada. De modo que,

$$f(c) \leq f(x_0), \quad \forall c \in C.$$

Lo que extendemos a $\mathbb{R}^n$.



Es decir, si C convexo no vacío, $C \in \mathbb{R}^n$ y $x_0 \notin C$, entonces $\exists f \in (\mathbb{R}^n)^\#$ tal que $f(c) \leq f(x_0)$, $\forall c \in C$.

Ahora, el objetivo es demostrar el teorema para una separación, y como no se tiene que jugar con puntos interiores, ni con ceros. Sea $C = A - B$, como son disjuntos $0 \notin C \subseteq \mathbb{R}^n$. Por lo tanto,

$$\exists f \in (\mathbb{R})^\# : f(c) \leq f(0) =, \quad \forall c \in C.$$

Lo que equivale a:

$$f(a - b) \leq 0 \quad \forall a \in A, \forall b \in B.$$

Utilizando la linealidad de nuevo,

$$f(a) \leq f(b) \quad \forall a \in A, \forall b \in B.$$

Lo que podemos tomar supremos e ínfimos quedándonos

$$f(a) \leq \sup\{f(a) : a \in A\} = r = \inf\{f(b) : b \in B\} \leq f(b).$$

Esto será verdad siempre que un conjunto este acotado. ■

En $\mathbb{R}^n$ no hace falta encontrar puntos interiores porque la convexidad nos da puntos interiores. Si trabajamos en dimensión infinita los conjuntos convexos no tienen puntos interiores.

- El siguiente corolario, no es un teorema sobre $\mathbb{R}^n$, es sobre el espacio

Teorema 2.8 **Teorema del hiperplano soporte.** Sea C convexo con $\mathring{C} \neq \emptyset$ y $x_0 \in \partial C$. Entonces, $\exists f \in E^{\#} : f(x_0) = \sup\{f(c) : c \in C\} \geq f(c)$, $\forall c \in \overline{C}$.

Demostración.- Sea $x_0 \in \mathring{C}$. Por el teorema de separación $\exists f \in E^{\#}$ tal que

$$f(c) \leq f(x_0), \quad \forall c \in \mathring{C}.$$

Lo que implica,

$$f(c) \leq f(x_0) \quad \forall c \in \overline{C}.$$

Por lo tanto,

$$\sup\{f(c) : c \in C\} \leq f(x_0).$$

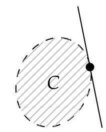
Como $x_0 \in \overline{C}$.. Entonces,

$$f(x_0) = \sup\{f(c) : c \in C\}.$$

Esto justifica el anterior teorema, donde el supremo o ínfimo tocarán a cualquiera de los dos conjuntos ajustándolos. ■

en general.

- La idea gráficamente sería:



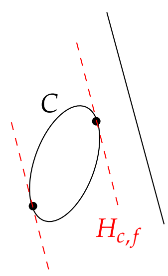
- Nuestro interés será encontrar una tangente que pase por el punto.
- Y al tomar la tangente, no me salgo del hiperplano, y el semiespacio donde queda el conjunto C.
- La tangente se ajusta al punto.

El hiperplano del gráfico de la derecha. Tiene un nombre especial, se llama **hiperplano soporte** que es lo que se definirá.

Definición 2.17 Dado un conjunto no vacío $C \subseteq E$, y $f \in E^{\#}$ al conjunto

$$H_{c,f} = \{x \in E : f(x) = \sup f(c) \text{ y } c \in C\}$$

se le llama **hiperplano soporte** a C en la dirección de f.

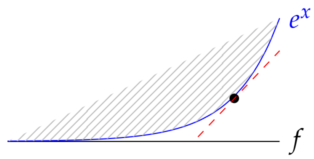


Si existe $x_0 \in C \cap H$, decimos que x_0 es un punto soporte.

Existe un saldo de la programación lineal a la programación convexa que en la programación lineal son importantes los puntos extremales, eso es porque los conjuntos son polígonos, donde puedo recuperar el polígono haciendo la envoltura convexa de los extremales. En optimización convexa, los sistemas son complejos por lo que el concepto extremal es menos importante y recobra más importancia el concepto de hiperplanos y los puntos soporte que serán el correspondiente a extremal, pero no tiene porque siempre tiene que ser extremal.

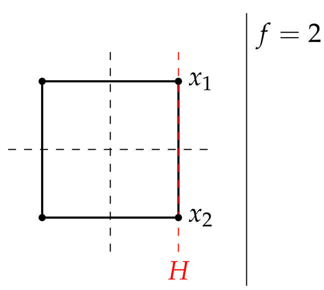
Ejemplo 2.6 Veamos un conjunto convexo cerrado pero que no tenga punto soporte. Si tomamos

$$C = \{(x, y) : y \geq e^x\}, \qquad f(x, y) = -y.$$



No hay puntos soporte. Pero lo interesante que por el teorema anterior siempre podremos construir un punto soporte. Ese soporte no tiene por que ser único.

Ejemplo 2.7 Veamos un conjunto convexo que tenga varios puntos soporte (simultaneamente).



$$C = \{(x, y) : \max\{x, y\} \leq 1\}$$

Si un hiperplano soporte H de un conjunto C tiene dos puntos soporte x_1, x_2. Entonces, todos los puntos del $[x_1, x_2]$ son puntos soporte de H.

- Si tenemos conjuntos cerrados, disjuntos y al menos uno de ellos es compacto. Entonces, puedo hacer separación fuerte.

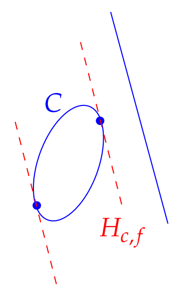
Teorema 2.9 **Hiperplano soporte para $\mathbb{R}^n$.** Sea $C \in \mathbb{R}^n$ no vacío y convexo. Sea $x_0 \in \partial C$. Entonces, existe un hiperplano soporte a C de forma que x_0 es un punto soporte. ■

Ejercicio 2.4 Sea $C \in \mathbb{R}^n$ Sea $C \in \mathbb{R}$ cerrado y convexo. Demuestra que C es la intersección de todos los semiespacios que lo contienen.

Demostración.- Primeramente notemos que, si cualquier semiespacio H tal que C está completamente dentro de H, entonces cada punto en C también pertenece a H. Ahora, si consideramos todos los posibles semiespacios que contienen a C, y tomamos su intersección, estamos buscando los puntos que son comunes a todos estos semiespacios. Dado que cada punto en C pertenece a cada uno de estos semiespacios (porque cada semiespacio contiene a C), cada punto en C también estará en la intersección de todos estos semiespacios. Por lo tanto, decimos que C está contenido en la intersección de todos los semiespacios que lo contienen.

Luego, sea un semiespacio un conjunto de la forma

$$H = \{x \in E : f(x) \leq \alpha\}, f \in E^{\#}, \alpha \in \mathbb{R}$$



y sea un punto x en la intersección de todos los semiespacios que contienen a C pero $x \notin C$. Como C es cerrado y convexo, y $x \notin C$, por el teorema de separación de Hahn-Banach, existe una función lineal f y un número real α tal que $f(x) > \alpha$ y $f(c) \leq \alpha$ para todo $c \in C$. Esto significa que x no está en el semiespacio

$$H = \{y \in \mathbb{R}^n : f(y) \leq \alpha\},$$

lo cual es una contradicción porque asumimos que x está en la intersección de todos los semiespacios que contienen a C. Por lo tanto, la intersección de todos los semiespacios que contienen a C debe estar contenida en C.

Así, demostramos que un conjunto cerrado y convexo C en $\mathbb{R}^n$ es la intersección de todos los semiespacios que lo contienen. ■

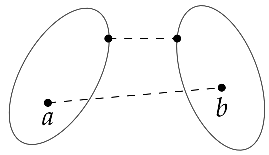
Teorema 2.10 Sean A convexo y compacto, B convexo y cerrado (con ejemplo de la exp que los dos sean cerrados no es suficiente), no vacíos. Tales que $A \cap B = \emptyset$. Entonces, existe $f \in (\mathbb{R}^n)^{\#}$. Tal que,

$$\sup f(a), a \in A < \inf f(b), b \in B.$$

Demostración.- Calculamos

$$d(A, B) = \in \{d(a, b) : a \in A, b \in B\}.$$

Es decir, tomo la distancia de un punto de uno y un punto del otro y tomamos el ínfimo. Pero si no hay compacidad, no están cerrados, esos puntos no existirán. De lo contrario se alanzarán.



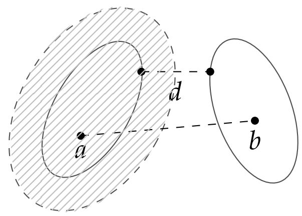
Lo que probaremos es que la $d(A,B) \neq 0$. Y al ser positiva nos permitirá encontrar un punto interior, por le cual utilizaremos el teorema de Hahn-Banach.

La distancia lo calcularemos con la Euclideana, porque estamos en $\mathbb{R}^n$; pero podríamos calcular para cualquier norma. Es decir,

$$d = d(A,B) = \inf\{\|a-b\|_2 : a \in A, b \in B\}.$$

(Ese ínfimo existirá, ya que está acotado por el cero inferiormente). Ahora, pueden pasar dos cosas:

- Si $d = 0$, tomamos una sucesión $\{a_n\}_n \subseteq A$ y $\{b_n\}_n \subseteq B$. De forma que,
  $$\|a_n - b_n\| \to d.$$
  El hecho de que las distancias entre los puntos se acerquen a algo, no quiere decir que los puntos se acerquen a nada. Pero como A es compacto,
  $$\{a_n\}_n \subseteq A$$
  tiene un sucesión convergente. Y ese límite de la sucesión, como el compacto es cerrado, también está en el compacto (cerrado por lo bordes y acotado que se puede dibujar),
  $$\left\{a_{n_k}\right\} \to a \in A. \qquad (1)$$
  Ya que,
  $$\|a_n - b_n\| \to 0.$$
  Por (1),
  $$\left\{b_{n_k}\right\} \to a.$$
  La demostración de $\left\{b_{n_k}\right\} \to a$ es: Si,
  $$\begin{aligned}\|b_{n_k} - a\| &\leq \|b_{n_k} - a_{n_k} + a_{n_k} - a\| \\ &\leq \|b_{n_k} - a_{n_k}\| + \|a_{n_k} - a\| \\ &\to 0 + 0 = 0.\end{aligned}$$
  Ahora, ya que B es cerrado, cualquier límite de cosas b están en B. Es decir,
  $$a \in A \cap B.$$
  Lo que es una contradicción, ya que $A \cap B = \emptyset$. Por lo tanto, $d \neq 0$.
- $d > 0$. Tomamos $A' = A + \frac{d}{2} B_{\mathbb{R}^n}$

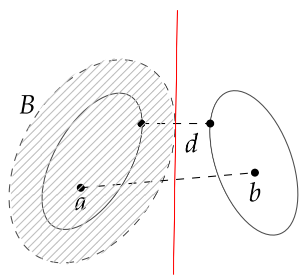


En particular, si $a \in A$. Entonces,

$$a + \frac{d}{2} B_{\mathbb{R}^n} \subseteq A'.$$

Esto implica que, a es un punto interior de A'. Luego, aplicamos el teorema de separación de los conjuntos, $\exists f : \mathbb{R}^n \to \mathbb{R}$ lineal, $f(a) < \inf f(b), b \in B, \ \forall a \in \operatorname{int} A'$. De donde,

$$f(a) < \inf f(b), b \in B, \ \forall a \in A.$$



Como A es compacto, existe

$$a_0 \in A : f(a_0) = \sup f(a), a \in A.$$

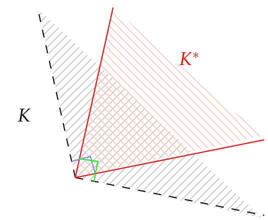
Entonces,

$$f(a) \geq f(a_0) < \inf \{f(b)\}, b \in B.$$

■

2.9 Conos duales

Los conos se establecen para ordenes en $\mathbb{R}^n$ que son multicriterio. Donde hablaremos de conos duales, que es importante ya que cuando hagamos optimización y nuestro dominio sea un cono, lo dualizaremos, cuando hagamos el dual del problema, el cono que nos saldrá es el cono dual. Entonces necesitamos saber como es el cono dual para poder trabajar con el cono dual.



Partimos de

$$K \in \mathbb{R}^n$$

que es un cono. Y a partir de aquí se define su cono dual como:

$$\begin{aligned} K^* &= \left\{ y \in \mathbb{R}^n : y^T x \geq 0, \forall x \in K \right\} \\ &= \left\{ y \in \mathbb{R}^n : \langle y, x \rangle \geq 0, \forall x \in K \right\}. \end{aligned}$$

Lo que tenemos que hacer es tomar lineas perperdiculares a las dos lineas que forman el cono, dando nos da un cono dual.

Ejemplo 2.8 Si $\|\cdot\|$ es una norma en $\mathbb{R}^n$. Entonces,

$$K = \{(x,t) \in \mathbb{R}^n : \|x\| \leq t\} \subseteq \mathbb{R}^{n+1}$$

(En realidad es cómo representar la gráfica de la función de la norma, y tomar el cono que se forma; es decir, una dimensión más).

Definición 2.18 La **norma dual** de $(\mathbb{R}^n, \|\cdot\|)$. Y si define como:

$$\|y\| = \sup \left\{ \langle y, x \rangle : x \in B_{\|\cdot\|} \right\}.$$

Tenemos una norma en espacio de Banach, en espacio normado. Cuando hacemos el dual, cuando vemos a $\mathbb{R}^n$ como aplicaciones lineales y continuas, entonces nos sale una norma nueva, esa norma se llama norma dual, que es importante.

La definición de cono dual se lleva bien con la definición de arriba y el resultado sería esta proposición:

Proposición 2.2 Si tomamos $\{(x,t) : \|x\| \leq t\}^* = \{(y,s) : \|y\|^* \leq s\}$.

Si utilizamos el cono $K = \{(x,t) \in \mathbb{R}^n : \|x\| \leq t\} \subseteq \mathbb{R}^{n+1}$, que **se usa mucho en optimización** y dualizamos lo que tendremos es el cono dual de la proposición 2.2 que no es más la que define a norma dual. Es decir, tomaríamos la norma, la dibujaríamos y aplicaríamos todo lo que está por encima.

Demostración.- Tomemos un punto de uno de los lados y demostraremos que ese punto ahí, si y sólo si está en el otro. Para ello, sea

$$(y,s) \in \{(y,s) : \|y\|^* \leq s\}.$$

De donde,

$$\|y\|^* \leq s.$$

Utilizando la definición 2.18 de norma dual se tiene,

$$\sup \left\{ \langle y, x \rangle : x \in B_{\|\cdot\|} \right\} \leq s.$$

Luego, por el producto escalar,

$$y^T x \leq s, \ \forall x \in B_{\|\cdot\|}.$$

El hecho de que x esté en la bola, me permitirá tomar los múltiplos de x. En otras palabras si yo multiplico x por algún número. Entonces,

$$y^T x \leq st, \ \forall x \in \mathbb{R}^n, \ \forall t \geq \|x\|.$$

Demostremos que $y^T x \leq s, \ \forall x \in B_{\|\cdot\|}$ es equivalente a $y^T x \leq s, \ \forall x \in \mathbb{R}^n$. Dado $x \in \mathbb{R}^n, t \geq \|x\|$. Entonces, $\frac{x}{t} \in B_{\mathbb{R}^n}$

Ahora, vemos que el espacio en $\mathbb{R}^n$ es simétrico. Es decir, para cualquier y, tenemos $-y$. Por lo que,

$$-y^T x \leq st, \ \forall x \in \mathbb{R}^n, \ \forall t \geq \|x\|.$$

Lo que es equivalente a decir que,

$$\begin{aligned} st + y * Tx &\geq 0, \quad \forall x \in \mathbb{R}^n, \ \forall t \geq \|x\| \\ \langle (y,s), (x,t) \rangle &\geq 0, \quad \forall (x,t) \in K. \end{aligned}$$

Por lo tanto por definición,

$$(y,s) \in K^*.$$

■

Estos conos duales tiene una serie de propiedades interesantes ya que nos permitirá hacer order.

Propiedad 2.4 Sea K un cono en $\mathbb{R}^n$.

1) K^* es un cono convexo y cerrado.
2) $K_1 \subseteq K_2 \Rightarrow K_2^* \subseteq K_1^*$.
3) $\mathring{R} \neq \emptyset \Rightarrow K^*$ apuntado o punteagudo.
4) $\overline{K}$ y es apuntado $\Rightarrow K^*$ tiene interior no vacío.
5) $K^{**} = \overline{\text{co}\, K}$.

- Apuntado significa que no puede tener un vector y su opuesto. El único vector opuesto es el cero. Si $y, -y \in K^* \Rightarrow y = 0$.
- Tener punto interior y ser apuntado, son propiedades duales.
- La 5) nos dice que si realizo la operación dos veces lo que consigo es cerrar.

- Por propio, se refiere a que podemos tener un orden. En otras palabras, que sea apuntado, que tuviese un punto interior, que sea convexo, etc.

Corolario 2.1 Si K es propio. Entonces, K^* es propio. ■

3

Funciones convexas