APUNTES DE CONVEXIDAD Y OPTIMIZACIÓN.

Incluye demostraciones propuestos en clases (en azul).

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

Christian Limbert Paredes Aguilera

2023

1

Convexidad y Optimización

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,

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

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 $D_i\left(Ax-b\right)$.

$$D_i\left(Ax-b\right) = D_i\left(x_1A^1 + x_2A^2 + \cdots + x_nA^n - b\right) = A^i.$$

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.

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

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

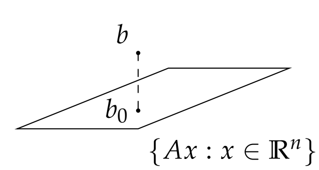
El cual es una ecuación normal.

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?:

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

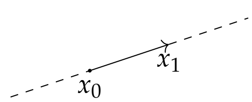
(Las ecuaciones normales vienen dadas por la perpendicularidad.)

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

El dominio serán conjuntos convexos o dominio efectivo.

Definición 1.1 **Lineal.**

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



- 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.

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

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

Definición 1.2 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).

- 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.

- 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 $x = y$.

Definición 1.3 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

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

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

Teorema 1.1 A es afín sii A contiene toda combinación afín de sus puntos.

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

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 1.2. *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 1.2 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 1.2 $A \subseteq \mathbb{R}^n$ 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 1.4 **Envoltura Afín.**

La envolura afín de B, $\text{Aff}(B)$, 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 1.5 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 1.2 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 1.6 **Topología de $\mathbb{R}^n$.**

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

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

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

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

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

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

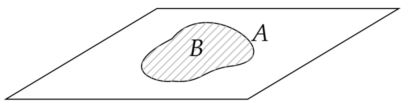
5) Se llama frontera de A, ∂A a la intersección $\overline{A} \cap \left(\overline{\mathbb{R}\backslash A}\right) = \overline{A} \backslash \text{int}(A)$ (Cualquier bola estará una parte en el interior y otra en el exterior del conjunto).

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.

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

6) Imaginamos un corte transversal para proyectar una imagen.

Ejemplo 1.3 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) $\text{relint}(A) = \emptyset$; $\text{relint}(B) = B \backslash A$.

Definición 1.7 **Combinación convexa.**

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

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 $\{x_1, \ldots, x_k\}$.

Observación 1.1 La definición para 2 puntos $\{x_1, x_2\}$ nos da las combinaciones convexas,

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

- En particular, las combinaciones convexas son los segmentos.

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.

Definición 1.8 **Convexo.**

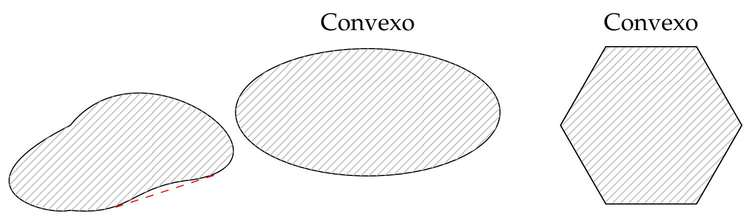
Un conjunto $C \in \mathbb{R}^n$ se dice convexo cuando C contiene las combinaciones convexas de sus puntos, si y sólo si

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

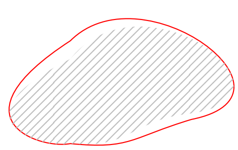
Un conjunto es convexo si dados dos puntos el segmento que los une se queda adentro.

- Decimos que C es cerrado para las combinaciones convexas. Es decir, no me salgo del conjunto.

Ejemplo 1.4



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



Definición 1.9 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 $\text{co}(A)$. También es equivalente a decir que

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

Ejercicio 1.1 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. ■

Definición 1.10 **Cono.**

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

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

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

Propiedad .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

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

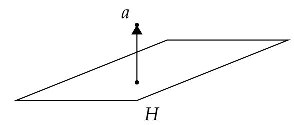
para $\lambda_i \geq 0$.

Definición 1.11 **Hiperplano.**

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

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

- El hiperplano es un caso particular del estudio convexo.
- Hiperplano:



Proposición 1.1 $H \subseteq \mathbb{R}^n$ es un hiperplano si y sólo si, H es un subespacio de dimensión $n-1$.

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

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

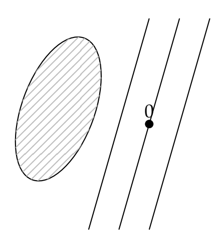
$$H = \left\{ x \in \mathbb{R}^n : a^T x = 0 \right\} = a^{\perp}$$

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:

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

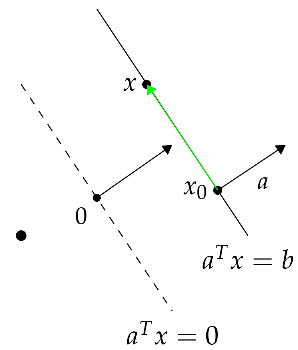
$$\left\{ x \in \mathbb{R}^n : a^T x > b \right\}$$

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.

- Un hiperplano en $\mathbb{R}^2$ será sencillamente las rectas.
- 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 1.5



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,

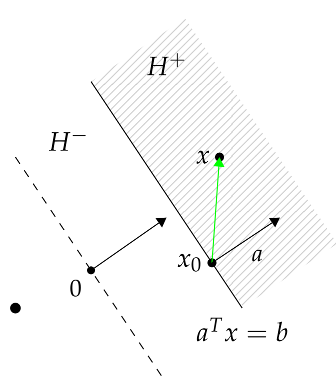
$$\langle (x - x_0), a \rangle = 0.$$

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 1.6



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

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

Luego,

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

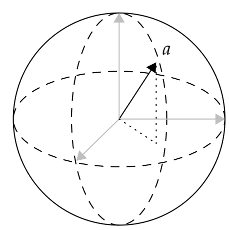
Por lo tanto,

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

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

1.3 Bolas Euclideas

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



Ejercicio 1.2 Demostrar que $B(c, r) = \left\{x \in \mathbb{R}^n : (x - c)^T (x - c) < r^2\right\}$ es convexo.

Demostración.- ■

Propiedad .2

$$B(c, r) = c + r$$