

APUNTES DE CONVEXIDAD Y OPTIMIZACIÓN.

Incluye demostraciones propuestos en clases ([en azul](#)).

Source: <https://github.com/soyfode/maticas/tree/master/investmat/src/convexOpt>

Christian Limbert Paredes Aguilera

2023

Introducción a la optimización convexa

1.1 Introducción

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

Donde,

$f_i : \mathbb{R}^n \rightarrow \mathbb{R}$
 f_0 : Función objetivo.
 f_j : Función Restricción donde $j = 1, \dots, m$.

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

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

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

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 $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ existirán algunas funciones el cual su dominio sera "tramoso".

Notación 1.1 Podemos escribir Ax como

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

$A^1 = A$ 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,

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

Nota 1.1. Imaginemos que tenemos

$$\begin{aligned} \{\min \quad & f(x) \\ \{\min \quad & f_0^2(x) \end{aligned}$$

- Si la 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 monótonas crecientes.
- Si el valor al cuadrado simplificará la solución, entonces podemos utilizarla. Esto nos permite que si no tenemos una función convexa podamos convexificarla.

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

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

El problema será una minimización global dada por: $\left\{ \begin{array}{l} \min : \|Ax - b\|_2^2 \\ \text{s.a.} \quad \emptyset. \end{array} \right.$

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 subíndice $_2$ significa la normal Euclidea. Que es la distancia normal que existe en \mathbb{R}^2 .

ello, encontraremos

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

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

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

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,

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

Observemos que,

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

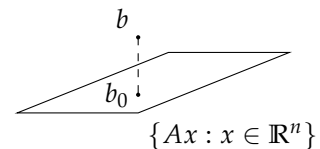
El cual es una ecuación normal.

Por último, veamos los argumentos geométricos. Notemos que,

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

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 algebraica-

mente?:

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

(Las ecuaciones normales vienen dadas por la perpendicularidad.)

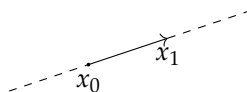
Conjuntos convexos

2.1 Conjuntos convexos de \mathbb{R}^n

El dominio serán conjuntos convexos o dominio efectivo.

Definición Lineal.
2.1

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



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

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

- 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.

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

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

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

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

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

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

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

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

$$\lambda_1 x_1 + \lambda_2 x_2, \quad \lambda_3 x_3 + \lambda_4 x_4, \quad \dots, \quad \lambda_{k-1} x_{k-1} + \lambda_k x_k, \quad \text{con } \sum_{i=1}^k y_i = 1.$$

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. ■

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

Nota 2.2. La definición de subespacio se refiere a tomar dos escalares y dos vectores, realizar la

Notación La suma de Minkowski es la operación de conjuntos; es decir, si $A, E \subseteq \mathbb{R}^n$. Entonces, *combinación lineal, donde esta combinación lineal no se saldrá del conjunto dado.*

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

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

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

$$\lambda, \mu \in \mathbb{R}, e_1, e_2 \in E \Rightarrow \lambda e_1 + \mu e_2 \in E.$$

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,

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

De donde,

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

Observemos que $\lambda_1 e_1 + \dots + \lambda_k e_k$ 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 Envoltura Afín.

2.4

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

2.5

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

Ejemplo 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

2.1

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

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 Topología de \mathbb{R}^n .

2.6

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

1) $a \in A$ está en el interior de A ($a \in \text{int}(A)$ o $a \in \overset{\circ}{A}$), 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 $A = \text{int}(A) = \overset{\circ}{A}$.

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

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

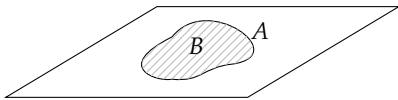
$$\{x \in \mathbb{R}^n : x \text{ está en el cierre de } A\}.$$

- 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 círculo 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) Se llama frontera de A , ∂A a la intersección $\overline{A} \cap (\overline{\mathbb{R} \setminus A}) = \overline{A} \setminus \text{int}(A)$ (Cualquier bola estará una parte en el interior y otra en el exterior del conjunto).
- 6) $a \in \text{relint}(A)$ si existe $\delta > 0$ tal que $B(a, \delta) \cap \text{Aff}(A) \subseteq A$.
- 5) Cualquier bola está en el interior cómo en el exterior del conjunto.
- 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 óptimo de una 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) $\overset{\circ}{A} = \emptyset$ | Si yo ponga una bola \mathbb{R}^3 , se saldrá del conjunto A .
- 4) $\overset{\circ}{B} = \emptyset$ | Ya que no existirá en el plano ninguna esfera.
- 5) $\text{relint}(A) = \emptyset$; $\text{relint}(B) = B \setminus A$.

Definición 2.7 Combinación convexa.

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

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

Al vector

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

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

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

- En particular, las combinaciones con-

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

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

vexas 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 2.8 **Convexo.**

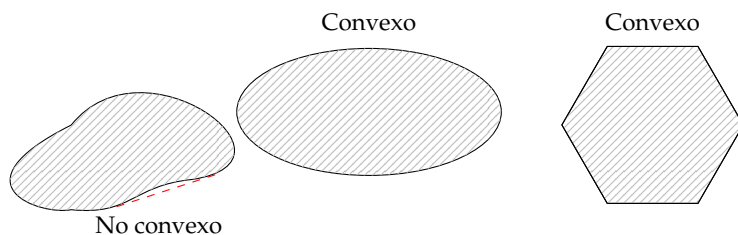
Un conjunto $C \in \mathbb{R}^n$ se dice convexo cuando C contiene las combinaciones convexas de sus puntos, si y sólo si

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

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

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

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

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

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 . Como 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 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.

- Un cono siempre contiene al origen.
- 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 .
- Un cono C es convexo si y sólo si

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

- C es un cono convexo si

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

para $\lambda_i \geq 0$.

Definición Hiperplano.

2.11

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

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

Proposición $H \subseteq \mathbb{R}^n$ 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

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

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:

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

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

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

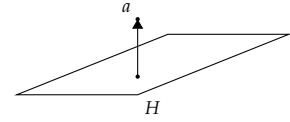
Esto, nos dará dos semiespacios dados por:

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

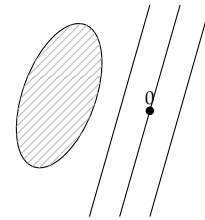
$$\{x \in \mathbb{R}^n : a^T x > b\}$$

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

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



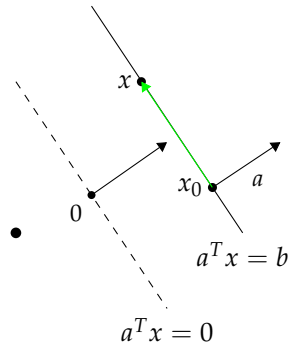
- En \mathbb{R} los hiperplanos son rectas.
- En \mathbb{R}^n , los hiperplanos serán uno menos de dimensión.



- 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.

Ejemplo
2.4

- La b nos dará una notación de distancia entre los hiperplanos.



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,

$$a^T (x - x_0) = 0$$

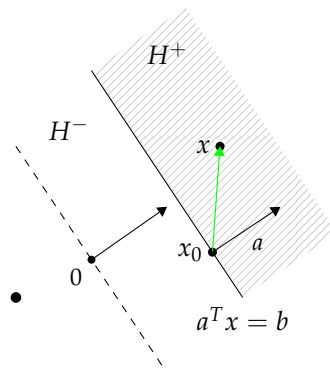
$$a^T x - a^T x_0 = 0$$

$$a^T x = a^T x_0$$

$$a^T x = b$$

Ejemplo
2.5

- Estas aplicaciones la llaman también aplicaciones del dual.



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,

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

Por lo tanto,

$$x \in H^+ \Leftrightarrow \langle (x - x_0), a \rangle = \cos [\text{ang}(x - x_0, a)] \|x - x_0\|_2 \|a\|_2 \geq 0$$

$$\Leftrightarrow \langle (x - x_0), a \rangle \geq 0$$

$$\Leftrightarrow \langle x, a \rangle \geq \langle x_0, a \rangle$$

$$\Leftrightarrow a^T x \geq a^T x_0$$

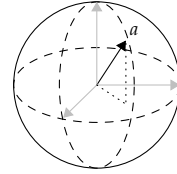
$$\Leftrightarrow a^T x \geq b$$

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

2.2 Bolas Euclideas

Tenemos que,

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



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

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

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,

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

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 2.12 Mediante la suma de Minkowsky 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

$$\mathcal{E} = \{x \in \mathbb{R}^n : (x - c)^T P^{-1} (x - c) \leq 1\} \quad (1)$$

$$= c + \{y \in \mathbb{R}^n : y^T P^{-1} y \leq 1\} \quad (y = x - c) \quad (2)$$

$$= c + \{y \in \mathbb{R}^n : y^T L D L^T y \leq 1\} \quad (3)$$

$$= c + \{Lz : z^T D z \leq 1\} \quad (z = L^T y) \quad (4)$$

$$= c + L \{z : z^T D z \leq 1\}. \quad (5)$$

donde

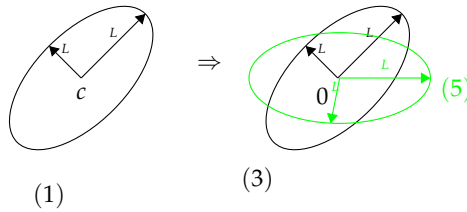
$$P = P^T > 0 \text{ (Simétrica y valores propios } > 0 \text{)}.$$

y

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

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 como producto de una matriz diagonal por dos matrices de cambio que en realidad son ortonormales.
- $z^T D z$, se puede hacer más grande o mas pequeña.
- Los vectores de L me dan los vectores que apuntan a la elipse.
- $z^T D z$ es el círculo.
- D serán las curvaturas principales de la elipse.
- La L gira la elipsoide.
- Los elipsoides se manejan para manejar imagenes donde incluye un objeto.



2.4 Bolas generales y conos asociados

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

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,

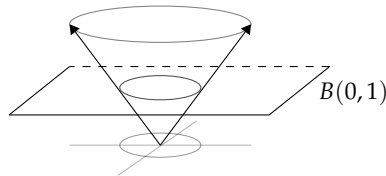
$$\begin{aligned} B(c, r) &= \{x \in \mathbb{R}^n : \|x - c\| \leq r\} \\ &= \left\{x \in \mathbb{R}^n : d_{\|\cdot\|}(x, c) \leq r\right\}. \end{aligned}$$

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

2.4.1 Cono asociado a una norma

Definición 2.13

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

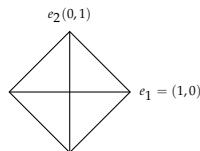


$$\{(x, t) : t = 1 \cap C_{\|\cdot\|}\} = \overline{B(x, t)}.$$

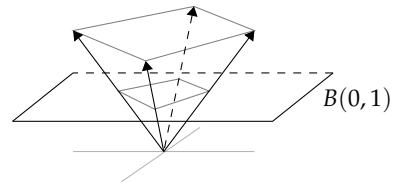
- Una vez que conoces la $B(0,1)$ se conoce todas las demás.
- No se puede derivar en el origen, ya que estará en el pico del cono.
- Podemos analizar todo lo que está dentro del cono,
- A esto se le llama epigrafo de la función, es decir dibujar la función y pintar todo lo que hay arriba.
- La definición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

Tipos de norma

- a) $\|\cdot\|_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, \|(x, y)\|_1 = |x| + |y|$. 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) $\|\cdot\|_\infty : \mathbb{R}^n \rightarrow \mathbb{R}$, $\|x, y\|_\infty = \max\{|x|, |y|\}$. Compara y coge la más grande.