1

Convexidad y Optimización

1.1 Introducción

$$(P)\begin{cases} \min & f_0(x) \\ & \\ s.a. & f_1(x) \leq b_1 \\ & f_2(x) \leq b_2 \\ & \vdots \\ & f_m(x) \leq b_m. \end{cases} \qquad \begin{array}{ll} f_i: & \mathbb{R}^n \to \mathbb{R} \\ f_0: & \text{Función objetivo.} \\ f_j: & \text{Función Restricción donde } j = 1, \ldots, m. \end{array}$$

- Las funciones objetivos en economía se les puede llamar función de coste.
- Las desigualdades tiene un truco, si multiplicamos por (-1) tenemos en la forma que decidamos.
- Maximizar es lo mismo que minimizar. Por lo que minimizaremos las funciones.

El objetivo de (P) es encontrar x^* el optimo ($\arg\min$) que cumple

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

Será en cualquier x que cumple las restricciones. Los puntos que no cumplen las condiciones no sirven para nada.

Al valor $f_0(x^*)$ se le llama valor optimo.

$f_i : \mathbb{R}^n \to \mathbb{R}$ Existirá algunas funciones que su dominio sera tranposo.

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

- Si los problemas son lineales se llama programación lineal.
- Cuando es convexa se llama optimización convexa.
- La habilidad es de identificar las restricciones y pasarlas a convexas.

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

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

Diremos que el un vector cualquiera sera vector columna.

Ahora, el problema será una minimización global dada por:

$$\begin{cases} \min : & \|A\mathbf{x} - b\|_2^2 \\ s.a. & \emptyset. \end{cases}$$

El subindice $_2$ significa la normal Euclidea. Que es la distancia normal que existe en $\mathbb{R}^2$.
El objetivo será encontrar la x donde la operación dada será la menor posible.
Nota 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 porque lo estoy elevando al cuadrado, pero el punto optimo lo será. Porque las funciones son monótomas crecientes. Si el valor al cuadrado me simplifica entonces podemos utilizarla. Esto nos permite que si no tengamos una función convexa podamos convexificarla.

Por diferenciabilidad:

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

Notación.- Podemos escribir Ax como

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

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,

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

Intentaremos demostrar el punto donde las parciales de $f_0 = 0$. Para 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 $D_i\left(Ax-b\right)$.

$$\begin{aligned} D_i\left(Ax-b\right) &= D_i\left(x_1A^1 + x_2A^2 + \cdots + x_nA^n - b\right) \\ &= A^i. \end{aligned}$$

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,

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

Observemos que,

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

En funciones convexas el extremo local será el mínimo global.

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

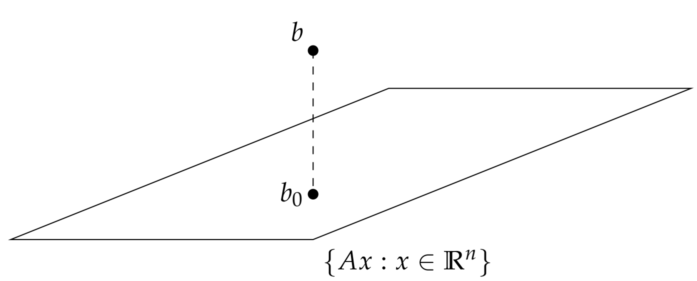
Que es una ecuación normal.

Argumentos geométricos

Sera el mismo calcular el mínimo de la distancia, calcular:

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

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



Si $b \in \{Ax : x \in \mathbb{R}^n\} \quad \Leftrightarrow \quad x^* \in \mathbb{R}^n : Ax^* = b$.
El valor optimo $f_0(x^*) = 0$.

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^* donde la solución es:

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

donde b_0 está en el plano, si estamos en $\mathbb{R}^3$

Cómo llegamos algebraicamente?

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

$\blacksquare$

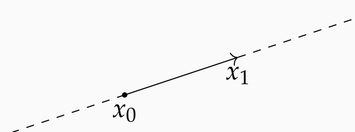
Las ecuaciones normales vienen por dada por la perpendicularidad.

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

El dominio tendrán que ser 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 me sale x_1 cuando valga cero me sale x_0 cuando es positivo va hacia la derecha y cuando es negativo va hacia la izquierda.
- Toda la recta nos da un concepto que denominamos Afín.
- Para la convexidad no es necesario tener la linea, solo necesitaremos un segmento.

$$\{x_0 + \lambda(x_1 - x_0) : \lambda \in [a, b]\}$$

- El segmento importante será el intervalo que denotaremos como:

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

  Es cualquier punto que este entre x_0 y x_1 del gráfico de arriba.

Definición 1.2 Sea $A \subseteq \mathbb{R}^n$. Se dice **Afín**, si para todo $x, y \in A$ se tiene que la $L(x, y) \subseteq A$. (Subespacios vectoriales desplazados).

- Un circulo no es afín ya que la linea es infinita y un circulo no.
- Un plano podría ser Afín
- La recta es afín
- Todo $\mathbb{R}^n$ es afín.
- Un único punto también es afín, dado que $x = y$.
- La diferencia entre espacio vectorial y espacio afín es que el espacio afín esta desplazada; es decir, no necesariamente pasa por el cero como en un subespacio vectorial.

Manejar el concepto de afín con lineas es un poco incomodo, entonces se utiliza el concepto de combinación afín

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

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

- Lo que decimos que es una combinación lineal de x_0 y x_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.

Demostración.- Ejercicio. ∎

- Quiere decir que este conjunto es estable para combinaciones lineales muy similar al concepto de subespacio vectorial.

Notación 1.1 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. ∎

Nota 1.1 La definición de subespacio es que tomo dos escalares y dos vectores, realizo la combinación lineal, y esta combinación lineal no se sale del conjunto dado.

Teorema 1.2 $A \in \mathbb{R}^n$ es afín sii existe un $E \subseteq \mathbb{R}^n$ subespacio vectorial 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, sabemos que $A \subseteq \mathbb{R}^n$ tal que $A = E + x_0$ para todo $x_0 \in A$ es afín. Entonces, E es un subespacio vectorial. Para demostrar que A es afín, probaré que cualquier combinación afín de elementos de A sigue estando en A. 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, $Aff(B)$, 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 $Affn$ 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.