"... Máma, la libertad siempre la llevarás
Dentro del corazón
Te pueden corromper, te puedes olvidar
Pero ella siempre está..."
Charly García

Curso 1.
Primera Reunión de Formas cuadráticas.

Definición

Dada una matriz $A \in \mathbb{R}^{n\times n}$ simétrica, una **forma cuadrática** en $\mathbb{R}^n$ es una función $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que: $Q(x) = x^T Ax$.

Ejemplo:

1. Si$A = \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}$, queda definida $Q(x) = [x_1\ x_2] \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

$$Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = [3x_1 + 2x_2\ \ 2x_1 + 5x_2]\,[x_1 x_2]$$

$$Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = (3x_1 + 2x_2)x_1 + (2x_1 + 5x_2)x_2 = 3x_1^2 + 2x_2x_1 + 2x_1x_2 + 5x_2^2$$

   Observemos que $a_{11} = 3$, $a_{22} = 5$ y $2a_{12} = 2a_{21} = 4$

$$Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = 3x_1^2 + 4x_1x_2 + 5x_2^2.$$

2. Si $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $Q(x) = 3x_1^2 - 2x_2^2 + 5x_3^2 + 6x_1x_2 - 4x_2x_3$.

   Inmediatamente podemos reconstruir la matriz $A \in \mathbb{R}^{3\times 3}$ que la define:

$$A = \begin{bmatrix} 3 & 3 & 0 \\ 3 & -2 & -2 \\ 0 & -2 & 5 \end{bmatrix} \text{ y } Q\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = [x_1\ \ x_2\ \ x_3] \begin{bmatrix} 3 & 3 & 0 \\ 3 & -2 & -2 \\ 0 & -2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Definición

Si: $Q(x) = x^T Ax,\ A \in \mathbb{R}^{2\times 2}$ simétrica, $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$, se llama **Curva de nivel** k al conjunto $C_k = \{x \in \mathbb{R}^2 : Q(x) = k,\ k \in \mathbb{R}\} \subset \mathbb{R}^2$.

Si: $Q(x) = x^T Ax,\ A \in \mathbb{R}^{3\times 3}$ simétrica, $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$, se llama **Superficie de nivel** k al conjunto $S_k = \{x \in \mathbb{R}^3 : Q(x) = k,\ k \in \mathbb{R}\} \subset \mathbb{R}^3$.

Es mucho más sencillo graficar este tipo de conjuntos cuando en la fórmula de la forma cuadrática no aparecen productos cruzados.

Cambio de variables.

Si $A \in \mathbb{R}^{n\times n}$ es una matriz simétrica, sabemos que A es diagonalizable ortogonalmente.

Existe una matriz $P \in \mathbb{R}^{n\times n}$ tal que $A = PDP^T$, con $D = diag(\lambda_1, \dots, \lambda_n)$ con λ_1 autovalor de $A\ \forall\ i = 1, \dots n$.

Entonces, $Q(x) = x^T Ax = x^T PDP^T x = (x^T P)D(P^T x) = (P^T x)^T D(P^T x)$.

Si hacemos el cambio de variables $y = P^T x$, obtenemos:

$x^T Ax = (x^T P)D(P^T x) = y^t Dy$

Entonces si $x,\ y \in \mathbb{R}^n :\ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ e $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

$$y^t Dy = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots \lambda_n y_n^2$$

Por lo tanto, $x^t Ax = k \Longleftrightarrow \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots \lambda_n y_n^2 = k$ con $y = P^T x$.

La principal ventaja de este cambio de variables es que elimina los productos cruzados y "mantiene la norma " de los vectores, pues si $y = P^T x$, como P es una matriz ortogonal, se cumple $||x|| = ||y||$.

ACLARACIÓN
Las formas cuadráticas se clasifican en positivas, negativas, semidefinidas positivas, semidefinidas negativas o indefinidas, según lo sean las matrices simétricas que las definen.

Optimización con restricciones

Vamos a estudiar cómo resolver problemas en los que necesitamos encontrar el máximo y mínimo de una forma cuadrática, $Q(x) = x^T Ax$, con una restricción en la norma de x.

Problemas de la forma :

$$\boxed{\text{Hallar } \max_{||x||^2=a^2} Q(x) \text{ y } \min_{||x||^2=a^2} Q(x)}$$

También podemos escribirlo con la notación:

$$\boxed{\text{Hallar } \max\{Q(x),\ \text{con } ||x||^2 = a^2\} \text{ y } \min\{Q(x),\ \text{con } ||x||^2 = a^2\}}$$

Cuanda $a = 1$ se suele decir que es el problema estándar de optimización de la formas cuadráticas.

Como siempre, empecemos por ver qué sucede en el caso más simple, por ejemplo cuando la matriz A, que define a la forma cuadrática, es una matriz diagonal.

Ejemplo:

Dada $Q(x) = 7x_1^2 - 3x_2^2 + 5x_3^2$, nos piden encontrar el máximo, el mínimo de esta función sujeto a la restricción $||x||^2 = 1$ y los puntos dónde se alcanza.

Resolución:

Busco el máximo de esta función cuando la calculo sobre todos los puntos de la esfera de radio 1:

$$7x_1^2 - 3x_2^2 + 5x_3^2 \leq 7x_1^2 + 7x_2^2 + 7x_3^2 \leq 7(x_1^2 + x_2^2 + x_3^2)$$

Entonces si $||x||^2 = 1$, $7x_1^2 - 3x_2^2 + 5x_3^2 \leq 7$, o sea 7 es un cota superior del conjunto de todos los resultados de aplicar la forma cuadrática sobre la esfera de radio 1.

Es obvio que si $x_M = \pm[1\ \ 0\ \ 0]^T \Rightarrow Q(x_M) = 7$. Entonces 7 es un valor del conjunto $\{Q(x) : ||x||^2 = 1\} \Rightarrow 7$ es máximo de este conjunto. Podría quedar la duda sobre si estos son los dos únicos puntos donde se alcanza este valor. ¿Cómo nos sacamos la duda? Simplemente planteando las igualdades convenientes.

Busco todos los $x \in \mathbb{R}^3$ tal que $||x||^2 = 1$ y $Q(x) = 7$, quedarían las ecuaciones:

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{1}$$
$$7x_1^2 - 3x_2^2 + 5x_3^2 = 7 \tag{2}$$

A la segunda ecuación le restamos la primera ecuación multiplicada por 7:

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &= 1 \\ -10x_2^2 - 2x_3^2 &= 0 \end{aligned}$$

Entonces, de la segunda ecuación obtenemos que $x_2 = x_3 = 0$ y reemplazando en la primera ecuación obtenemos $x_1^2 = 1 \Rightarrow x_1 = \pm 1$.

Entonces los únicos puntos de la esfera dónde se alcanza el máximo de la función son $x_M = \pm[1 \;\; 0 \;\; 0]^T$

Lo mismo vamos a hacer para calcular el mínimo valor de Q sobre la esfera unitaria:

$$7x_1^2 - 3x_2^2 + 5x_3^2 \geq -3x_1^2 - 3x_2^2 - 3x_3^2 \geq -3\underbrace{(x_1^2 + x_2^2 + x_3^2)}_{=1} = -3$$

Entonces si $||x||^2 = 1$, $7x_1^2 - 3x_2^2 + 5x_3^2 \geq -3$, o sea -3 es un cota inferior del conjunto de todos los resultados de aplicar la forma cuadrática sobre la esfera de radio 1.

Es obvio que si $x_{min} = \pm[0 \;\; 1 \;\; 0]^T \Rightarrow Q(x_{min}) = -3$.
Entonces, $-3 \in \{Q(x) : ||x||^2 = 1\} \Rightarrow -3$ es mínimo de este conjunto pues es una cota inferior y pertenece al conjunto.
Para ver si hay otros puntos, planteamos:

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{3}$$
$$7x_1^2 - 3x_2^2 + 5x_3^2 = -3 \tag{4}$$

A la segunda ecuación le restamos la primera ecuación multiplicada por -3:

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &= 1 \\ 10x_1^2 + 8x_3^2 &= 0 \end{aligned}$$

De la segunda ecuación obtenemos que $x_1 = x_3 = 0$ y reemplazando en la primera ecuación obtenemos $x_2^2 = 1 \Rightarrow x_2 = \pm 1$.

Entonces los únicos puntos de la esfera dónde se alcanza el mínimo de la función son $x_m = \pm[0 \;\; 1 \;\; 0]^T$.

Otro Ejemplo super sencillo:

Veamos que pasa si alguno de los coeficientes de la diagonal se repite:

Buscamos máximo de $Q(x) = 7x_1^2 + 7x_2^2 + 2x_3^2$ sujeto a la restricción $||x||^2 = 1$

Resolución:

Otra vez, para buscar el máximo acotamos a derecha la función, reemplazando todos los coeficientes por el máximo:

$$7x_1^2 + 7x_2^2 + 2x_3^2 \leq 7x_1^2 + 7x_2^2 + 7x_3^2 = 7\underbrace{(x_1^2 + x_2^2 + x_3^2)}_{=1} = 7$$

El máximo entonces es 7 pues si buscamos los puntos donde se alcanza este máximo, planteamos:

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{5}$$
$$7x_1^2 + 7x_2^2 + 2x_3^2 = 7 \tag{6}$$

A la segunda ecuación le restamos la primera ecuación multiplicada por 7:

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &= 1 \\ -5x_3^2 &= 0 \end{aligned}$$

Entonces, $x_3 = 0$ y reemplazando en la otra ecuación obtenemos:
$x_1^2 + x_2^2 = 1$

Por lo que todos los puntos donde se alcanza ese máximo son los puntos de la circunferencia definida por esa ecuación. (Son los puntos que resultan de la intersección del plano de autovectores asociados a $\lambda = 7$ y la esfera de radio 1.)

Tarea para el hogar hallar el mínimo de Q con la restricción $||x||^2 = 1$

Después de estos ejemplos vemos que es sencillo resolver el problema cuando la matriz que define la función cuadrática es una matriz diagonal.

Entonces, como la matriz que define a una forma cuadrática es una matriz simétrica podemos solucionar el problema fácilmente con un cambio de variables ortogonal, como el que hicimos para sacar los productos cruzados.
$A = PDP^T \Rightarrow Q(x) = x^TPDP^Tx \Rightarrow$ si $y = P^Tx, x^TPDP^Tx = y^TDy$.

Además $||x|| = ||y||$ porque P es ortogonal.

Por lo tanto, el problema estándar de optimización $\max\limits_{||x||^2=1} x^TAx$ y $\min\limits_{||x||^2=1} x^TAx$ se transforma con este cambio de variable en:

$\max\limits_{||y||^2=1} y^TDy$ y $\min\limits_{||y||^2=1} y^TDy$.
Con este cambio de variables se demuestra el Teorema de Rayleigh.

Resolvamos el problema de hallar el máximo de x^TAx para una matriz simétrica A.
Buscamos $\max\limits_{||x||^2=1} x^TAx$, haciendo el cambio de variables ortogonal $y = P^Tx$, sabemos que

debemos resolver:

$\max\limits_{||y||^2=1} y^T Dy$, con $\lambda_1 \ldots, \lambda_n$, autovalores de A.

Vamos a ordenar los autovalores de mayor a menor: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. De esta manera, $\lambda_{MAX} = \text{máx}\{\lambda_i : \lambda_i \text{autovalor de} A\} = \lambda_1$ y $\lambda_m = \text{mín}\{\lambda_i : \lambda_i \text{autovalor de} A\} = \lambda_n$.

Además, si $P = [v_1| \ldots |v_n]$ es la matriz ortogonal formada por los autovectores de A y $\dim(S_{\lambda_1})= k$, entonces $\{v_1, \ \ldots, v_k\}$ es una BON de $S_{\lambda=\lambda_1}$ y $\lambda_1 = \cdots = \lambda_k$.

$$y^T Dy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \leq \lambda_1 y_1^2 + \cdots + \lambda_1 y_n^2 = \lambda_1 (y_1^2 + \cdots + y_n^2) = \lambda_1 \underbrace{||y||^2}_{=1} = \lambda_1 = \lambda_{MAX}.$$

Al igual que en los casos que vimos en la introducción, es evidente que λ_{MAX} no es sólo una cota superior, pues si tomamos $y_M = [1 \ 0 \ \ldots 0]^T$ obtenemos que $y_M^T \ D \ y_M = \lambda_1$. Entonces λ_1 es el máximo valor de $y^T Dy$.

Busquemos ahora, los puntos de norma 1 que cumplen $y^T Dy = \lambda_1$:

$$y_1^2 + \cdots + y_k^2 + y_{k+1}^2 + \cdots + y_n^2 = 1 \tag{7}$$
$$\lambda_1 y_1^2 + \cdots + \lambda_1 y_k^2 + \lambda_{k+1} y_{k+1}^2 + \cdots + \lambda_n y_n^2 = \lambda_1 \tag{8}$$

A la segunda ecuación le restamos la primera ecuación multiplicada por λ_1:

$$\begin{aligned} y_1^2 + \cdots + y_k^2 + \cdots + y_n^2 &= 1 \\ \underbrace{(\lambda_{k+1} - \lambda_1)}_{<0} y_{k+1}^2 + \cdots + \underbrace{(\lambda_n - \lambda_1)}_{<0} y_n^2 &= 0 \end{aligned}$$

entonces, de la última ecuación sale que $y_{k+1} = \cdots = y_n = 0$ y reemplazando en la primera ecuación:

$y_1^2 + \cdots + y_k^2 = 1.$

Entonces, en las nuevas variables obtuvimos:

$\boxed{\max\limits_{||y||^2=1} y^T Dy = \lambda_1 = \lambda_{MAX}}.$

Los puntos donde se alcanza este extremo son de la forma:

$\boxed{y_M = [y_1 \ldots \ y_k \ 0 \ldots 0]^T \text{ con } y_1^2 + \cdots + y_k^2 = 1.}$

Volviendo a las variables originales podemos concluir que :

$$\max_{||x||^2=1} x^T Ax = \lambda_{MAX}, \text{ con } \lambda_{MAX} \text{máximo autovalor de } A.$$

$$\text{Además como } y = P^T x \Rightarrow x_M = Py_M = [v_1|\dots|v_k|v_{k+1}|\dots v_n] \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Entonces $x_M = y_1 v_1 + \cdots + y_k v_k$ con $y_1^2 + \cdots + y_k^2 = 1$.
Encontramos el máximo de $x^T Ax$ y los puntos donde se alcanza:

$$\boxed{\max_{||x||^2=1} x^T Ax = \lambda_{MAX}.}$$

$$\boxed{x_M = y_1 v_1 + \cdots + y_k v_k \text{ con } y_1^2 + \cdots + y_k^2 = 1.}$$

Antes de enunciar el Teorema de Rayleigh, recordemos la igualdad (a):

$$y^T Dy \leq \lambda_1 ||y||^2.$$

Podemos escribir entonces:

$$x^T Ax = y^T Dy \leq \lambda_1 ||y||^2 = \lambda_1 ||x||^2$$

$\forall$ A **simétrica** se cumple:

$$\boxed{x^T Ax \leq \lambda_{MAX} ||x||^2 \text{ y } x^T Ax = \lambda_{MAX} ||x||^2 \text{ si } x \in S_{\lambda_{MAX}}.}$$

Otra forma en la que puede enunciarse el mismo resultado es:

$$\boxed{\tfrac{x^T Ax}{||x||^2} \leq \lambda_{MAX}, \forall x \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\} \text{ y } \tfrac{x^T Ax}{||x||^2} = \lambda_{MAX} \text{ si } x \in S_{\lambda_{MAX}} - \{0_{\mathbb{R}^n}\}.}$$

Hemos demostrado en definitiva, el Teorema de Rayleigh:

Teorema de Rayleigh: Si $A \in \mathbb{R}^{n\times n}$ es una matriz simétrica con λ_{MAX} y λ_m los autovalores máximo y mínimo de A respectivamente y $S_{\lambda_{MAX}}$ y S_{λ_m} los respectivos autoespacios, entonces:

- $\lambda m\ ||x||^2 \leq x^T Ax \leq \lambda_{MAX}||x||^2$, $\forall x \in \mathbb{R}^n$ **Desigualdad de Rayleigh**.
  La igualdad se cumple en los respectivos autoespacios.
- $\max\limits_{x\neq 0} \frac{x^T Ax}{||x||^2} = \lambda_{MAX}$, se alcanza en $x \in \mathbb{S}_{\lambda_{MAX}} - \{0_{\mathbb{R}^n}\}$.
- $\min\limits_{x\neq 0} \frac{x^T Ax}{||x||^2} = \lambda_m$, se alcanza en $x \in \mathbb{S}_{\lambda_m} - \{0_{\mathbb{R}^n}\}$.

En consecuencia, gracias a la primera desigualdad, tenemos tresueltos infinitos problemas de optimización:

- $\max\limits_{||x||^2=a^2} x^T Ax = \lambda_{MAX}a^2$ y $x_{MAX} \in S_{\lambda_{MAX}} \cap \{||x||^2 = a^2\}$
- $\min\limits_{||x||^2=a^2} x^T Ax = \lambda_m a^2$ y $x_m \in S_{\lambda_m} \cap \{||x||^2 = a^2\}$

Ejemplo:

Dada A=$\begin{bmatrix} 5 & -1 & 0 \\ -1 & 5 & 0 \\ 0 & 0 & 6 \end{bmatrix}$, si $Q(x) = x^T Ax$, se pide hallar máximo y mínimo de $Q(x)$ sujeto a la restricción $||x|| = 1$

Resolución:

Ya sabemos por el Teorema de Railegh que $\lambda m||x||^2 \leq x^T Ax \leq \lambda_{MAX}||x||^2$, $\forall x \in \mathbb{R}^n$

Así que buscamos autovalores de A:

$$P_A(\lambda) = \begin{vmatrix} (\lambda-5) & 1 & 0 \\ 1 & (\lambda-5) & 0 \\ 0 & 0 & (\lambda-6) \end{vmatrix} = (\lambda-6)[(\lambda-5)^2-1] = 0 \Leftrightarrow \lambda = 6 \text{ o } (\lambda-5)^2 = 1$$

Obtenemos que las raíces son $\lambda_1 = 6$ de multiplicidad 2 y $\lambda = 4$ simple.
Entonces, en este caso $\lambda_{MAX} = 6$ y $\lambda_m = 4$
Los valores máximo y mínimo de la forma cuadrática con la restricción $||x||^2 = 1$ ya los conocemos:

$$\max_{||x||^2=1} Q(x) = 6 \text{ y } \min_{||x||^2=1} Q(x) = 4$$

Busquemos ahora los puntos donde se alcanza la igualdad, para eso necesitamos los respectivos autoespacios:

$$S_{\lambda=6} = \text{Nul}(6I - A) = \text{Nul}\left(\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{gen}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

$$S_{\lambda=4} = \text{gen}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Si buscamos bases ortonormales de cada autoespacio, tenemos:

$$\left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\} \text{ base de } S_{\lambda=6}.$$

$$\left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}\right\} \text{base de } S_{\lambda=4}.$$

El máximo se alcanza en el conjunto $x \in S_{\lambda=6} \cap \{||x||^2 = 1\}$:

$x = \alpha \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, tal que $||x||^2 = 1$.Como los vectores son ortogonales y de norma 1, obtenemos:

$$x_{MAX} = \alpha \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ con } \alpha^2 + \beta^2 = 1.$$

El mínimo se alcanza en el conjunto $x \in S_{\lambda=4} \cap \{||x||^2 = 1\}$:

$$x_m = \pm \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$