"... ...Desprejuiciados son los que vendrán,
y los que están ya no me importan más
los carceleros de la humanidad
no me atraparán,
dos veces con la misma red..."
Charly García

2ª Reunión formas cuadráticas . Curso 1.
Últimos comentarios sobre optimización de Formas Cuadráticas.

Siempre trabajamos con $A \in \mathbb{C}^{n\times n}$ o $A \in \mathbb{R}^{n\times n}$ y consideramos el P.I canónico en $\mathbb{C}^n$ ($\langle x, y\rangle = \overline{y}^T x$) o $\mathbb{R}^n$ ($\langle x, y\rangle = y^T x$).

Recordemos:

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: $Q(x) = x^T Ax$.

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

Cambio de variables.

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

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

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

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

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

Entonces $x^t Ax = c \iff \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots \lambda_n y_n^2 = c$ 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 $||x|| = ||y||$.

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

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

Con este teorema entonces, tenemos resueltos **todos los problemas de optimización de la forma:**

Hallar $\max\limits_{||x||^2=a^2} Q(x)$ y $\min\limits_{||x||^2=a^2} Q(x)$ y los puntos donde se alcanzan.

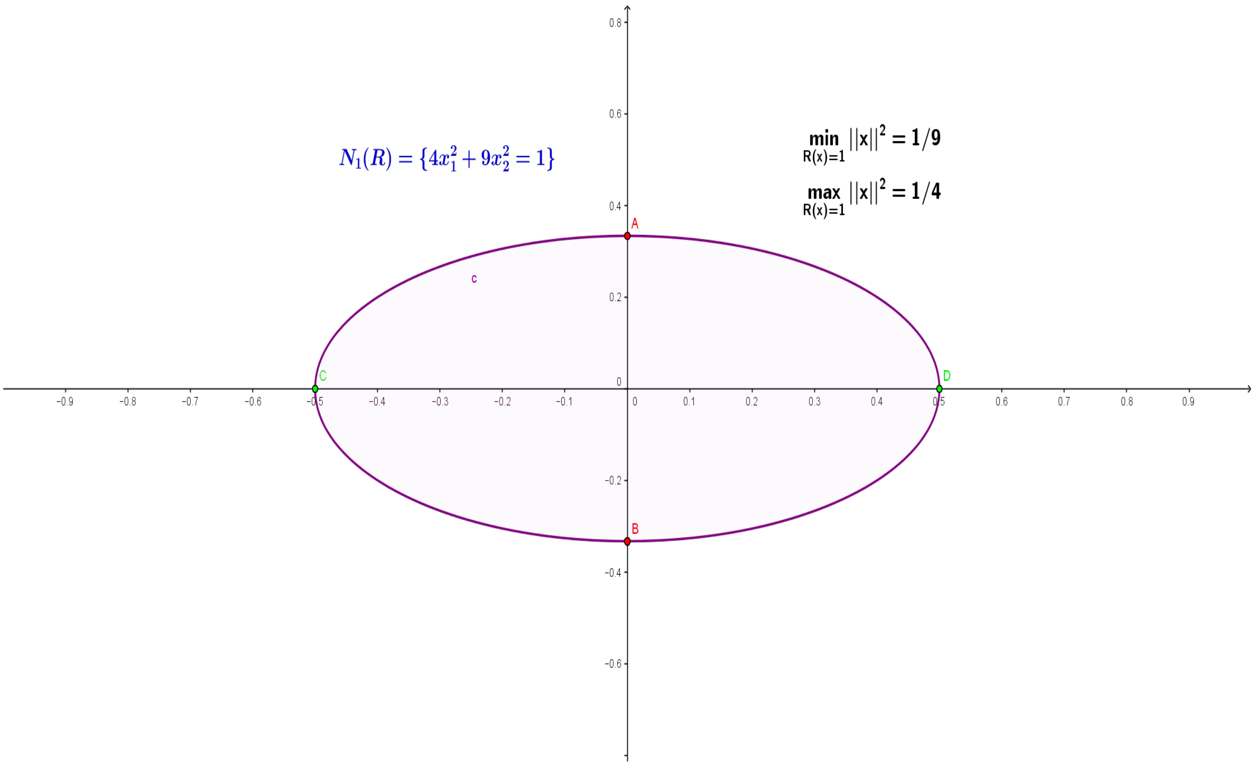
Para toda $Q(x) = x^T Ax$, con $A \in \mathbb{R}^{n\times n}$, simétrica.

Ahora nos vamos a preocupar de otros problemas de optimización.
Ejemplo:

1. Dada $R(x) = 4x_1^2 + 9x_2^2$,

   a)Grafique la curva de nivel $R(x) = 1$ y encuentre $\max\{\| x \|^2 \colon R(x) = 1\}$ y $\min\{\| x \|^2 \colon R(x) = 1\}$
   b) Demuestre analíticamente lo averiguado en el punto anterior.

   Resolución:



b) La resolución analítica de este problema se desprende del teorema de Rayleigh:

$$\lambda_{min}\, ||x||^2 \leq 4x_1^2 + 9x_2^2 \leq \lambda_{MAX}||x||^2,\ \forall x \in \mathbb{R}^2$$

$$4\, ||x||^2 \leq 1 \leq 9||x||^2,\ \forall x \in \mathbb{R}^2$$

De aquí entonces tenemos que:

$\min\{\| x \|^2 \colon R(x) = 1\} = \frac{1}{9}$ y se alcanza en $x_{min} = \pm[0\ \ \frac{1}{3}]^T$.

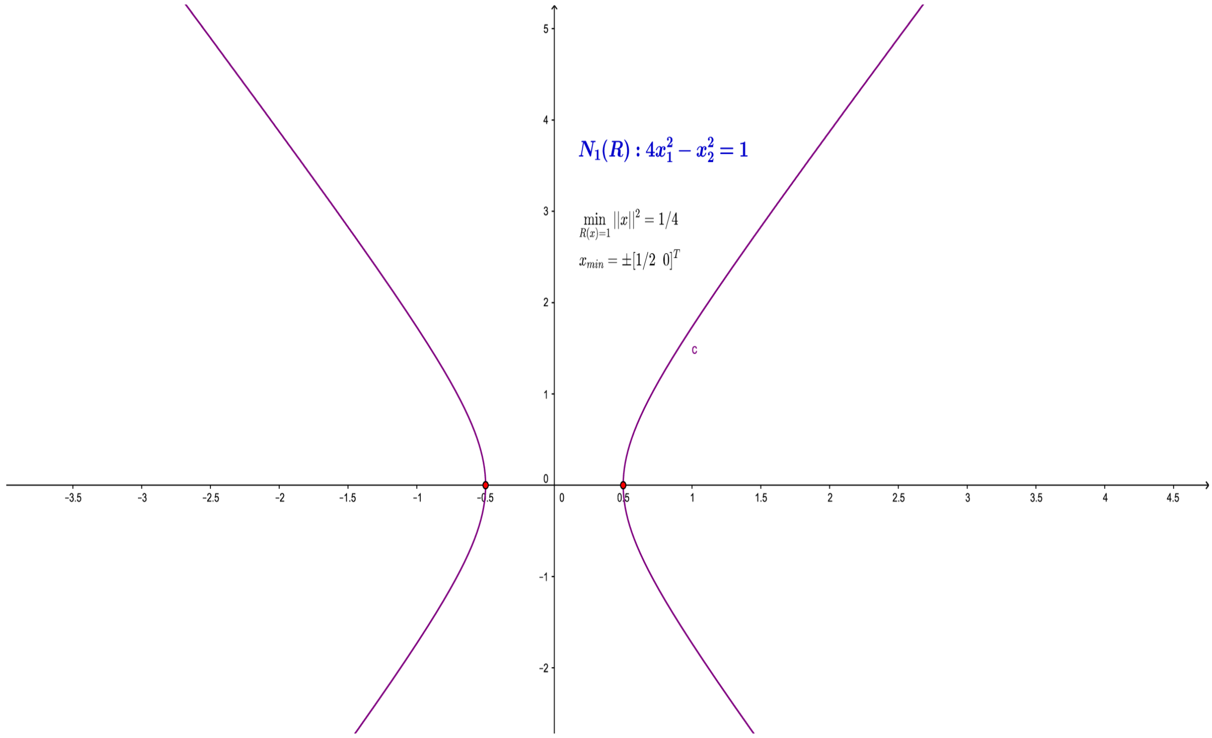
$\max\{\| x \|^2 \colon R(x) = 1\} = \frac{1}{4}$ y se alcanza en $x_{MAX} = \pm[\frac{1}{2}\ \ 0]^T$.

2. Dada $R(x) = 4x_1^2 - x_2^2$.

a)Grafique la curva de nivel $R(x) = 1$ y compruebe que mín $\{\| x \|^2 : R(x) = 1\} = 1/4$, pero no existe máx $\{\| x \|^2 : R(x) = 1\}$.

b) Demuestre analíticamente lo averiguado en el punto anterior.

Resolución:



b) Para demostrar que existe el mínimo del conjunto basta con usar la desigualdad de Rayleigh:

$$\lambda min \ ||x||^2 \leq R(x) \leq \lambda_{MAX} ||x||^2, \ \forall x \in \mathbb{R}^2$$

$$-1 \ ||x||^2 \leq R(x) \leq 4||x||^2, \ \forall x \in \mathbb{R}^2$$

Cuando $R(x) = 1$:

$$-1 \ ||x||^2 \leq 1 \leq 4||x||^2, \ \forall x \in \{x / R(x) = 1\}$$

Por lo tanto, del lado derecho de la desigualdad obtenemos:

$$\frac{1}{4} \leq ||x||^2 \ \forall x \in \{x/R(x) = 1\}$$

Entonces gracias al Teorema de Rayleigh podemos asegurar que :

$$\min_{4x_1^2 - x_2^2 = 1} ||x||^2 = \frac{1}{4} \text{ y este valor se alcanza en } x_{min} = \pm[1/2 \ \ 0]^T$$

En el gráfico vemos claramente que los puntos en esta curva no están acotados. Analíticamente esto se puede demostrar exhibiendo una sucesión de puntos en la curva de nivel cuya norma tiende a infinito.

Por ejemplo, busco puntos, $P_n = [x_{1n}x_{2n}]^T$ tales que :

$$4x_{1n}^2 - x_{2n}^2 = 1 \text{ y } ||P_n||^2 = x_{1n}^2 + x_{2n}^2 = n^2$$

Haciendo cuentas encontramos una sucesión que cumple lo pedido: $x_{1n} = \sqrt{\frac{n^2+1}{5}}$ y $x_{2n} = \sqrt{\frac{4n^2+1}{5} - 1}$, $n \in \mathbb{N}$

Para poder generalizar, veamos *tomados de la mano* algunas características de las formas cuadráticas que nos van a servir para solucionar algunos problemas.

Observaciones:

En todo lo que sigue $Q(x) = x^T Ax$ con $A \in \mathbb{R}^{n\times n}$ simétrica, y λ_{MAX} y λ_{min} los autovalores máximo y mínimo de A respectivamente.

a $Q(\alpha x) = \alpha^2 Q(x)$

b Si $\lambda_1 \neq \lambda_2$ son autovalores de A con v_1 y v_2 autovectores asociados respectivos de $A \Rightarrow$ $Q(\alpha v_1 + \beta v_2) = \alpha^2 Q(v_1) + \beta^2 Q(v_2) = \alpha^2 \lambda_1 ||v_1||^2 + \beta^2 \lambda_2 ||v_2||^2$

c Si $\lambda_{min} < \lambda_{MAX} \Rightarrow Q(x)$ toma todos los valores del intervalo $(\lambda_{min},\ \lambda_{MAX})$ sobre la circunferencia de radio 1, $(||x|| = 1)$.

d Si Q es una forma cuadrática definida positiva $\Rightarrow$ sus conjuntos de nivel c¿0, están acotados $(N_c(Q) == \{x \in \mathbb{R}^n : Q(x) = c, c > 0\})$.

Demostración:

e Si Q es una forma cuadrática definida positiva $\Rightarrow Q(x) = x^T Ax$, con A simétrica definida positiva $\Rightarrow$ todos los autovalores de A son estrictamente positivos $\Rightarrow 0 < \lambda_{min} \leq \lambda_{MAX}$. Por el Teorema de Rayleigh:

$$\lambda min\ ||x||^2 \leq x^T Ax \leq \lambda_{MAX} ||x||^2,\ \forall x \in \mathbb{R}^n$$

.

Sobre el conjunto de nivel $N_c(Q)$ se cumple $Q(x) = x^T Ax = c$:

$$\lambda min\ ||x||^2 \leq c \leq \lambda_{MAX} ||x||^2,\ \forall x \in N_c(Q)$$

Como $\lambda min > 0 \Rightarrow ||x||^2 \leq \frac{c}{\lambda_{min}}$, esto quiero decir que el conjunto está acotado.

f Si $Q(x)$ es una forma cuadrática semidefinida o indefinida $\Rightarrow$ sus conjuntos de nivel no están acótados. si la función cuadrática es semidefinida o indefinida tiene por lo menos un autovalor nulo o un autovalor negativo.

Si la forma cuadrática es semidefinida positiva el menor autovalor de A es $\lambda_{min} = 0$ y $\lambda_{MAX} > 0$ es fácil ver que en ese caso el conjunto de nivel $x^T Ax = c$ no está acotado. Sea v_1 un autovector de A asociado a λ_{min} y sea w un autovector de A asociado a λ_{MAX}, podemos suponer v y w vectores unitarios.
Consideremos la sucesión de puntos de la forma:

$$v_n = nv_1 + \sqrt{\frac{c}{\lambda_{MAX}}}\, w$$

Por la observación b:

$$Q(v_n) = Q(nv_1 + \sqrt{\frac{c}{\lambda_{MAX}}}\, w) = n^2 Q(v_1) + \sqrt{\frac{c}{\lambda_{MAX}}} Q(w) = n^2 \cdot 0 + \lambda_{MAX} \cdot \lambda_{MAX} = c$$

O sea, $v_n \in N_c(Q)$ y $||v_n||^2 = ||nv_1 + \sqrt{\frac{c}{\lambda_{MAX}}}\, w||^2 = n^2 + \frac{c}{\lambda_{MAX}}$

Por lo tanto $\lim\limits_{n \to +\infty} v_n = +\infty$

Si la forma cuadrática es indefinida $\Rightarrow \lambda_{min} < 0$ y $\lambda_{MAX} > 0$, existen vectores unitarios v_1 autovector de A asociado a λ_{min} y w autovector de A asociado a $\lambda_{MAX} > 0$. Siempre se puede construir una sucesión de puntos de manera tal que ,$P_n = \alpha v_1 + \beta w$ cumplan a la vez:

$$Q(P_n) = \alpha^2 \lambda_{min} + \beta^2 \lambda_{MAX} = c \text{ y } ||P_n||^2 = \alpha^2 + \beta^2 = n^2 \to +\infty$$

Sólo hace falta verificar que las dos ecuaciones que quedan planteadas para α y β tiene solución. Por esto se puede afirmar:

El conjunto de nivel $N_c(Q)$ **no está acotado** si la forma cuadrática es indefinida o semidefinida.

g Si Q es una forma cuadrática definida positiva es muy fácil resolver el problema de optimización:

$$\boxed{\max_{Q(x)=c} ||x||^2 \text{ y } \min_{Q(x)=c} ||x||^2 \text{ y los puntos donde se alcanzan.}}$$

Otra vez, por la desigualdad de Rayleigh:

$$\lambda min\, ||x||^2 \leq x^T Ax \leq \lambda_{MAX} ||x||^2,\ \forall x \in \mathbb{R}^n$$

.

Sobre el conjunto de nivel $N_c(Q)$ se cumple $Q(x) = x^T Ax = c$:

$$\lambda min\, ||x||^2 \leq c \leq \lambda_{MAX} ||x||^2,\ \forall x \in N_c(Q)\ \textbf{(a)}$$

Entonces, como $\lambda min > 0$ de la desigualdad izquierda obtenemos:

$$||x||^2 \leq \frac{c}{\lambda_{min}}$$

y la igualdad se cumple sobre el subespacio asociado, entonces:

$$\boxed{\max_{Q(x)=c} ||x||^2 = \tfrac{c}{\lambda_{min}} \text{y se alcanza en } x \in S_{\lambda_{min}} \cap \{Q(x) = c\} = S_{\lambda_{min}} \cap \left\{||x||^2 = \tfrac{c}{\lambda_{min}}\right\}}$$

Trabajando con el lado derecho de la desigualdad **(a)**, obtenemos:

$$\frac{c}{\lambda_{MAX}} \leq ||x||^2$$

Entonces:

$$\boxed{\min_{Q(x)=c} ||x||^2 = \tfrac{c}{\lambda_{MAX}} \text{y se alcanza en } x \in S_{\lambda_{MAX}} \cap \{Q(x) = c\} = S_{\lambda_{MAX}} \cap \{||x||^2 = \tfrac{c}{\lambda_{MAX}}\}}$$

Por lo visto en el item f. cuando estamos trabajando con una forma cuadrática, Q, **indefinida o semidefinida**, existe el $\min\limits_{Q(x)=c} ||x||^2 = \frac{c}{\lambda_{MAX}}$ pero no existe el $\max\limits_{Q(x)=c} ||x||^2$, pues el conjunto de nivel no está acotado.

Si enfrentamos el problema más general :

$$\boxed{\text{Hallar } \max_{R(x)=1} Q(x) \text{ y } \min_{R(x)=1} Q(x) \text{ y los puntos donde se alcanzan.}}$$

donde $R(x)$ y $Q(x)$ son formas cuadráticas.

Se puede asegurar la existencia de los dos extremos cuando R es definida positiva, o sea $R(x) = x^T Bx,\ B$ matriz simétrica definida positiva.

Si B es una matriz definida positiva puede realizarse un cambio de variables, para llevar el problema a un problema estandar.

Si B es definida positiva, existen $P ortogonal y \mathrm{D}_1 = diag(\beta_1, \ldots, \beta_n)$, con $\beta_i > 0$, $\forall i = 1, \ldots, n$ tales que $B = PD_1P^T$.
Además, si llamamos $D^* = diag(\sqrt{\beta_1}, \ldots, \sqrt{\beta_n})$, podemos escribir $B = PD^*D^*P^T$, por lo que la restricción $R(x) = x^T Bx = x^T PD^*D^*P^T x = 1$, puede escribirse mediante el cambio de variables $z = D^*P^T x$, de la forma: $z^T z = 1$.

Si realizamos el cambio de coordenadas sobre la cuadrática $Q(x) = x^T Ax$, debemos reemplazar la variable x en función de z. Si $z = D^*P^T x$, entonces $x = PD^{*-1}z$, con $D^{*-1} = diag(\frac{1}{\sqrt{\beta_1}}, \ldots, \frac{1}{\sqrt{\beta_n}})$. $Q(x) = x^T Ax = (PD^{*-1}z)^T APD^{*-1}z = z^T D^{*-1}P^T APD^{*-1}z$.
La matriz $C = D^{*-1}P^T APD^{*-1}$, resulta ser una matriz simétrica cuando A lo es. Por lo tanto con este cambio de variables hemos llevado el problema original a un problema estandar en las nuevas variables. Y resolveremos

$$\boxed{\text{Hallar } \max_{||z||^2=1} z^T Cz \text{ y } \min_{||z||^2=1} z^T Cz \text{ y los puntos donde se alcanzan.}}$$

Una vez resuelto el problema en las nuevas variables, para señalar los puntos donde se alcanzan los extremos en las variables originales, ”volvemos” con el cambio de variables calculando $x = PD^{*-1}z$.