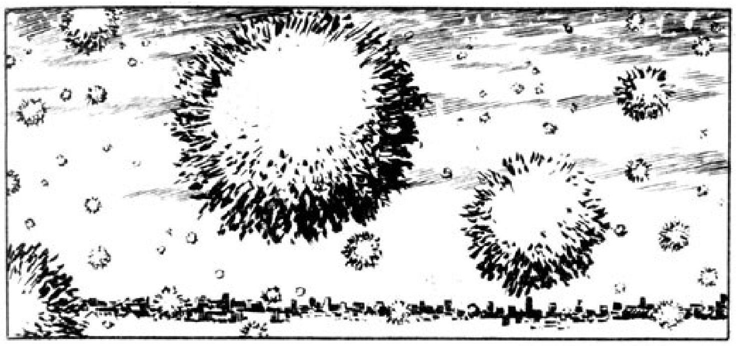
Álgebra II (Curso 23)
Segundo cuatrimestre, 2021

NOTAS EN LA EMERGENCIA SANITARIA:

BORRADORES PARA LA CLASE DEL 8 DE DICIEMBRE

Sebastian GRYNBERG



El único héroe válido es el héroe "en grupo",
nunca el héroe individual, el héroe solo.

H. G. OESTERHELD

ÍNDICE

1. Formas cuadráticas

1.1. Presentación.

Nomenclatura. En todo lo que sigue $\mathrm{Sim}_n(\mathbb{R})$ es el conjunto de todas las matrices simétricas de $\mathbb{R}^{n\times n}$. Decir que $A \in \mathrm{Sim}_n(\mathbb{R})$ significa que $A \in \mathbb{R}^{n\times n}$ y que $A^T = A$.

Definición 1.1. *Una* forma cuadrática *en* $\mathbb{R}^n$ *es una función* $Q : \mathbb{R}^n \to \mathbb{R}$ *que admite la siguiente representación*

$$Q(x) = x^T A x, \tag{1}$$

con $A \in \mathrm{Sim}_n(\mathbb{R})$.

Nota Bene. Nótese que si $A \in \mathrm{Sim}_n(\mathbb{R})$, entonces

$$\begin{aligned} x^T A x &= \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij} x_i x_j \end{aligned} \tag{2}$$

Procedimiento. Si una forma cuadrática de $\mathbb{R}^n$ se presenta en la forma

$$Q(x) = \sum_{i=1}^{n} c_i x_i^2 + \sum_{1\leq i<j\leq n} c_{ij} x_i x_j \tag{3}$$

y se desea representarla en la forma $Q(x) = x^T A x$, con $A \in \mathrm{Sim}_n(\mathbb{R})$, se puede proceder de la siguiente forma. Se definen

- $a_{ii} := c_i =$ el coeficiente de x_i^2, para todo $i \in \{1, \dots n\}$
- $a_{ij} := \frac{1}{2} c_{ij} =$ la mitad del coeficiente del término cruzado $x_i x_j$,

y se recurre a la simetría de A para completar los coeficientes restantes. Con esas definiciones la expresión dada en (3) adopta la forma

$$Q(x) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij} x_i x_j = x^T A x \tag{4}$$

con $A = (a_{ij}) \in \mathrm{Sim}_n(\mathbb{R})$.

Catecismo matemático.

Pregunta: ¿Qué es la matriz de una forma cuadrática?

Respuesta: La matriz de una forma cuadrática $Q : \mathbb{R}^n \to \mathbb{R}$ es la única matriz simétrica $A \in \mathbb{R}^{n\times n}$ tal que $Q(x) = x^T A x$ para todo $x \in \mathbb{R}^n$.

1.2. Conjuntos de nivel.

Dado $c \in \mathbb{R}$ el conjunto de todas las soluciones de la ecuación $Q(x) = c$ se denomina *el conjunto de nivel* c de Q y lo denotaremos mediante

$$\mathcal{N}_c(Q) := \{x \in \mathbb{R}^n : Q(x) = c\}.$$

1.3. Diagonalización: ejes principales.

Cambios de variables. Sea $M \in \mathbb{R}^{n\times n}$ una matriz inversible. Nótese que si se realiza el cambio de variables $x = My$ la expresión de la forma cuadrática $Q(x) = x^T Ax$, en las nuevas variables y, adopta la forma

$$\tilde{Q}(y) = Q(My) = y^T M^T AMy.$$

El siguiente Teorema muestra que la naturaleza de una forma cuadrática $Q(x) = x^T Ax$, con $A \in \mathrm{Sim}_n(\mathbb{R})$, está determinada por los autovalores de la matriz A.

Teorema 1.2 (Teorema de los ejes principales)**.** *Sea $A \in \mathrm{Sim}_n(\mathbb{R})$. Existe un cambio de variables ortogonal $x = Py$ que transforma la forma cuadrática $Q(x) = x^T Ax$ en una forma cuadrática de la forma $\tilde{Q}(y) = y^T \Lambda y$ sin términos cruzados.*

Demostración. El *teorema espectral* para matrices simétricas garantiza la existencia de una matriz ortogonal P y una matriz diagonal $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ tales que $A = P\Lambda P^T$. Efectuando el cambio de variables $x = Py$ se obtiene que

$$x^T Ax = y^T P^T APy = y^T \Lambda y.$$

En este nuevo sistema de coordenadas la forma cuadrática considerada adopta la forma

$$\tilde{Q}(y) = y^T \Lambda y = \sum_{i=1}^{n} \lambda_j y_j^2, \tag{5}$$

que, como se puede observar a simple vista, no tiene términos cruzados. □

Nota Bene. Nótese que si $P = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$ e $y \in \mathbb{R}^n$, la expresión (5) se puede escribir en la forma

$$Q(Py) = Q\left(\sum_{j=1}^{n} y_j u_j\right) = \sum_{j=1}^{n} \lambda_j y_j^2. \tag{6}$$

Nomenclatura. Las columnas de la matriz P que aparece en el teorema anterior se denominan *ejes principales* de la forma cuadrática Q.

Nota Bene. Nótese que el efecto de la diagonalización ortogonal de la matriz A es producir una rotación del sistema de coordenadas canónico para que, en el nuevo sistema de coordenadas, el gráfico del conjunto de nivel $x^T Ax = c$ se vuelva evidente. Por ejemplo, si todos los autovalores de $A \in \mathbb{R}^{3\times 3}$ son positivos, la expresión (5) deja claro que el gráfico de la superficie de nivel $x^T Ax = c$, para $c > 0$, es un elipsoide centrado en el origen.

1.4. Formas canónicas.

Formas canónicas. Existe una matriz inversible $M \in \mathbb{R}^{n\times n}$ tal que el cambio de variables $x = My$ permite expresar a Q en la forma

$$\tilde{Q}(y) = y^T \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_r \end{bmatrix} y, \tag{7}$$

donde p, q y r son, respectivamente, las cantidades de autovalores positivos, negativos y nulos de A, (contados con sus multiplicidades). La representación de Q en la forma (7) se denomina *la forma canónica de Q*.

Demostración. Como $A \in \mathrm{Sim}_n(\mathbb{R})$, existe una diagonalización $A = P\Lambda P^T$ con $P \in \mathbb{R}^{n\times n}$ ortogonal y $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n\times n}$, donde $\lambda_1 \geq \cdots \geq \lambda_p > 0$, $\lambda_{p+1} \leq \cdots \leq \lambda_{p+q} < 0$, y $\lambda_{p+q+1} = \cdots = \lambda_n = 0$. Definiendo $D \in \mathbb{R}^{n\times n}$ mediante

$$D = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_p}, \sqrt{-\lambda_{p+1}}, \dots, \sqrt{-\lambda_{p+q}}, 1, \dots, 1),$$

resulta que $M = PD^{-1}$ satisface lo pedido. □

1.5. Cátalogo de superficies de nivel.

Se trata de un breve recordatorio de objetos geométricos que presuponemos conocidos.

1. Sea $Q(x)$ la forma cuadrática en $\mathbb{R}^3$ definida por
$$Q(x) := x_1^2 + x_2^2 + x_3^2.$$
Sus conjuntos de nivel son
$$\mathcal{N}_c(Q) = \begin{cases} \text{Esfera centrada en el origen de radio } \sqrt{c} & \text{si } c > 0, \\ \{0_{\mathbb{R}^3}\} & \text{si } c = 0, \\ \emptyset & \text{si } c < 0. \end{cases}$$
2. Sea $Q(x)$ la forma cuadrática en $\mathbb{R}^3$ definida por
$$Q(x) := x_1^2 + x_2^2 - x_3^2.$$
Sus conjuntos de nivel son
$$\mathcal{N}_c(Q) = \begin{cases} \text{Hiperboloide de una hoja} & \text{si } c > 0, \\ \text{Cono} & \text{si } c = 0, \\ \text{Hiperboloide de dos hojas} & \text{si } c < 0. \end{cases}$$
Se trata de tres superficies de revolución alrededor del eje x_3: la primera y la tercera se obtienen rotando la hipérbola $x_1^2 - x_3^2 = c$, y la segunda rotando la recta $x_1 = x_3$.
3. Sea $Q(x)$ la forma cuadrática en $\mathbb{R}^3$ definida por
$$Q(x) := x_1^2 + x_2^2.$$
Sus conjuntos de nivel son
$$\mathcal{N}_c(Q) = \begin{cases} \text{Cilindro circular} & \text{si } c > 0, \\ \text{Recta } x_1 = x_2 = 0 & \text{si } c = 0, \\ \emptyset & \text{si } c < 0. \end{cases}$$
4. Sea $Q(x)$ la forma cuadrática en $\mathbb{R}^3$ definida por
$$Q(x) := x_1^2 - x_2^2.$$
Sus conjuntos de nivel son
$$\mathcal{N}_c(Q) = \begin{cases} \text{Cilindro hiperbólico} & \text{si } c \neq 0, \\ \text{Par de planos que se cortan} & \text{si } c = 0. \end{cases}$$

5. Sea $Q(x)$ la forma cuadrática en $\mathbb{R}^3$ definida por

$$Q(x) := x_1^2.$$

Sus conjuntos de nivel son

$$\mathcal{N}_c(Q) = \begin{cases} \text{Par de planos paralelos} & \text{si } c > 0, \\ \text{Par de planos coincidentes} & \text{si } c = 0, \\ \emptyset & \text{si } c < 0. \end{cases}$$

1.6. Clasificación.

Sea $Q(x) = x^T Ax$ una forma cuadrática con matriz $A \in \mathrm{Sim}_n(\mathbb{R})$. Se dice que Q (o que su matriz simétrica A) es

(a) *definida positiva* si $Q(x) > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$,
(b) *semidefinida positiva* si $Q(x) \geq 0$ para todo $x \in \mathbb{R}^n$ y existe $x_0 \neq 0$ tal que $Q(x_0) = 0$.
(c) *definida negativa* si $Q(x) < 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$,
(d) *semidefinida negativa* si $Q(x) \leq 0$ para todo $x \in \mathbb{R}^n$ y existe $x_0 \neq 0$ tal que $Q(x_0) = 0$.
(d) *indefinida* si $Q(x_1) > 0$ para algún $x_1 \in \mathbb{R}^n$ y $Q(x_2) < 0$ para algún $x_2 \in \mathbb{R}^n$.

Proposición 1.3. *Sea $Q(x) = x^T Ax$ una forma cuadrática con matriz $A \in \mathrm{Sim}_n(\mathbb{R})$. Entonces*

(a) *Q es definida positiva si y solamente si $\sigma(A) \subset (0, +\infty)$.*
(b) *Q es semidefinida positiva si y solamente si $\sigma(A) \subset [0, +\infty)$ y $0 \in \sigma(A)$.*
(c) *Q es definida negativa si y solamente si $\sigma(A) \subset (-\infty, 0)$.*
(d) *Q es semidefinida negativa si y solamente si $\sigma(A) \subset (+\infty, 0]$ y $0 \in \sigma(A)$.*
(e) *Q es indefinida si y solamente si $\sigma(A) \cap (-\infty, 0) \neq \emptyset$ y $\sigma(A) \cap (0, \infty) \neq \emptyset$.*

Demostración. Se realiza utilizando el *teorema de los ejes principales* y las definiciones cada una de las propiedades de Q enunciadas en la proposición. □

Nota Bene. Nótese que por definición, vale que

1. si Q es una forma cuadrática definida positiva, entonces 0 es el único punto de $\mathbb{R}^n$ en el que se realiza el mínimo global de $Q(x)$;
2. si Q es una forma cuadrática definida negativa, entonces 0 es el único punto de $\mathbb{R}^n$ en el que se realiza el máximo global de $Q(x)$.

Comentario. Este tipo de resultados forman la base sobre la que se construyen los criterios de clasificación de los puntos críticos de un campo escalar $f : \mathbb{R}^n \to \mathbb{R}$ de clase C^2 que se estudian en Análisis Matemático II. Allí se muestra que si $x_0 \in \mathbb{R}^n$ es un punto crítico de f, en el sentido de que $\nabla f(x_0) = 0$, usando el desarrollo de Taylor de f centrado en x_0, resulta que

$$f(x_0 + h) = f(x_0) + \frac{1}{2} h^T \mathcal{H}_f(x_0) h + o(\|h\|^2),$$

donde $H_f(x_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)$ es la matriz Hessiana de f evaluada en x_0, esa expresión permite deducir que, por ejemplo, x_0 es un mínimo local cuando la forma cuadrática $Q(h) = h^T \mathcal{H}_f(x_0) h$ es definida positiva.

1.7. Imagen de los autoespacios.

Consideramos una forma cuadrática de $\mathbb{R}^n$ definida por $Q(x) = x^T Ax$, con $A \in \mathrm{Sim}_n(\mathbb{R})$. Queremos saber cómo actúa Q sobre los autoespacios de A

$$\mathbb{S}_\lambda = \mathrm{nul}(A - \lambda I), \quad \lambda \in \sigma(A).$$

Si $\mu(\lambda) = \dim(\mathbb{S}_\lambda)$ es la multiplicidad de λ, consideramos una diagonalización ortogonal de A de la forma $A = P\Lambda P^T$, con

$$P = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{R}^{n\times n} \;\text{ y }\; \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n),$$

donde los primeros $\mu(\lambda)$ elementos de la diagonal coinciden con λ y las primeras $\mu(\lambda)$ columnas de P son una base ortonormal de $\mathbb{S}_\lambda$.

Decir que $x \in \mathbb{S}_\lambda$ y escribir que $x = Py$ significa que $x = \sum_{j=1}^{\mu(\lambda)} y_j u_j$, y entonces

$$Q(x) = Q(Py) = \sum_{i=1}^{\mu(\lambda)} \lambda y_j^2 = \lambda \sum_{j=1}^{\mu(\lambda)} y_j^2 = \lambda\|y\|^2 = \lambda\|x\|^2.$$

Esto es así porque $\|y\| = \|Py\| = \|x\|$ y además $y_j = 0$ para todo $j > \mu(\lambda)$.

Lema 1.4. *Sea $Q(x) = x^T Ax$, con $A \in \mathrm{Sim}_n(\mathbb{R})$, una forma cuadrática en $\mathbb{R}^n$. Sea $\sigma(A)$ el espectro de A y para cada $\lambda \in \sigma(A)$ sea $\mathbb{S}_\lambda = \mathrm{nul}(A - \lambda I)$ el autoespacio correspondiente al autovalor λ de A. Vale que*

$$Q(x) = \lambda\|x\|^2 \quad \textit{para todo} \quad x \in \mathbb{S}_\lambda.$$

El resultado anterior nos permite caracterizar geométricamente la acción de Q sobre $\mathbb{R}^n$. Como

$$\mathbb{R}^n = \bigoplus_{\lambda\in\sigma(A)} \mathbb{S}_\lambda, \quad \text{con } \mathbb{S}_\lambda \perp \mathbb{S}_{\lambda'} \text{ para cada } \lambda \neq \lambda',$$

todo $x \in \mathbb{R}^n$ se descompone univocamente en la forma

$$x = \sum_{\lambda\in\sigma(A)} x_\lambda, \quad \text{con } x_\lambda \in \mathbb{S}_\lambda, \text{ para cada } \lambda \in \sigma(A)$$

Utilizando (6) tenemos que

$$Q(x) = \sum_{\lambda\in\sigma(A)} \lambda\|x_\lambda\|^2. \tag{8}$$

En palabras, *la imagen de x por Q se obtiene superponiendo los cuadrados de las longitudes de las componentes de x sobre cada uno de los autoespacios de la matriz de Q, cada uno multiplicado por su respectivo autovalor.*

1.8. Distancias al origen.

El resultado anterior permite analizar la distancia al origen de los puntos del conjunto de nivel $\mathcal{N}_c(Q)$. En primer lugar,

$$Q(x) = \lambda_m\|x\|^2 \text{ para todo } x \in \mathrm{nul}(A - \lambda_m I),$$

$$Q(x) = \lambda_M\|x\|^2 \text{ para todo } x \in \mathrm{nul}(A - \lambda_M I),$$

donde $\lambda_m = \min \sigma(A)$ y $\lambda_M = \max \sigma(A)$. Si $x_\lambda \neq 0$ para algún $\lambda \notin \{\lambda_m, \lambda_M\}$, la mutua ortogonalidad de los autoespecios de A y el Teorema de Pitágoras permiten deducir que

$$\lambda_m \|x\|^2 < Q(x) < \lambda_M \|x\|^2 \text{ para todo } x \notin \mathrm{nul}(A - \lambda_m I) \cup \mathrm{nul}(A - \lambda_M I).$$

Forma definida positiva. En lo que sigue examinaremos el caso en que la forma cuadrática es definida positiva. Los otros casos se analizan de la misma manera y quedan a cargo del lector.

Decir que $Q(x) = x^T A x$ es definida positiva significa que $Q(x) > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$, esto equivale a decir que todos los autovalores de A son positivos. Esto tiene las siguientes consecuencias, por un lado los conjuntos de nivel $\mathcal{N}_c(Q)$ se comportan de la siguiente manera

$$\mathcal{N}_c(Q) = \begin{cases} \text{Elipsoide centrado en el origen} & \text{si } c > 0, \\ \{0_{\mathbb{R}^n}\} & \text{si } c = 0, \\ \emptyset & \text{si } c < 0. \end{cases}$$

Por otro lado, $0 < \lambda_m \leq \lambda_M$. Los casos no triviales se producen cuando $c > 0$. En el caso en que $\lambda_m = \lambda_M$, el elipsoide resultante coincide con la esfera centrada en el origen cuyo radio es $\sqrt{\frac{c}{\lambda_m}}$ y todos sus puntos están a la misma distancia del origen. Queda por analizar qué pasa cuando $c > 0$ y $\lambda_m < \lambda_M$. En este caso tenemos que

$$\begin{aligned} c &= \lambda_m \|x\|^2 \text{ para todo } x \in \mathrm{nul}(A - \lambda_m I) \cap \mathcal{N}_c(Q), \\ c &= \lambda_M \|x\|^2 \text{ para todo } x \in \mathrm{nul}(A - \lambda_M I) \cap \mathcal{N}_c(Q), \\ \lambda_m \|x\|^2 < c &< \lambda_M \|x\|^2 \text{ para todo } x \in \mathcal{N}_c(Q) : x \notin \mathrm{nul}(A - \lambda_m I) \cup \mathrm{nul}(A - \lambda_M I). \end{aligned}$$

De allí se deduce que

1. Los puntos del conjunto de nivel $\mathcal{N}_c(Q)$ cuya distancia al origen es máxima son aquellos que también pertenecen al autoespacio de A correspondiente al autovalor mínimo λ_m. Además la distancia de esos puntos al origen es $\sqrt{\frac{c}{\lambda_m}}$.
2. Los puntos del conjunto de nivel $\mathcal{N}_c(Q)$ cuya distancia al origen es mínima son aquellos que también pertenecen al autoespacio de A correspondiente al autovalor máximo λ_M. Además la distancia de esos puntos al origen es $\sqrt{\frac{c}{\lambda_M}}$.