Mi camino a la perturbación espectral de matrices

ALAMA-GAMM/ANLA'2014, Barcelona

Juan-Miguel Gracia[1]

[1] *Department of Applied Mathematics, Statistics and O.R., University of the Basque Country UPV/EHU, Faculty of Pharmacy, Vitoria-Gasteiz, Spain*
E-mail: `juanmiguel.gracia@ehu.es`

Resumen

Trataré de narrar la ilación entre mis intereses matemáticos en el devenir del tiempo.

Como dijera Bellman, menos mal que el producto de matrices *no era conmutativo*: $AB \neq BA$. Todo lo demás se nos dio por añadidura. En el principio de la teoría de perturbación estuvieron los argumentos por continuidad. Me llamó poderosamente la atención un teorema que decía que tan cerca como quisiéramos de una matriz cuadrada compleja M existían matrices M' que tenían todos sus valores propios simples. Es decir, que el conjunto de tales matrices era *genérico*. Su demostración mediante la *resultante de Sylvester* $R(p,q)$ de dos polinomios $p(\lambda), q(\lambda)$ me abrió los ojos. Un cachivache —$R(p,q)$— que me había parecido una reliquia decimonónica la primera vez que la estudié. Ya en aquel tiempo vi con claridad que toda información concerniente a las raíces de un polinomio $p(\lambda)$ estaba encerrada en sus coeficientes. Sólo era cuestión de hacerlos *cantar*. Me gustaban mucho los lugares donde el álgebra se mezclaba con la topología y con el análisis matemático. En particular, el hecho de que los valores propios de una matriz dependieran con continuidad de los elementos a_{ij} de la matriz A me parecía sumamente interesante; aunque las *funciones* que expresaban la dependencia de valores propios de los elementos a_{ij} dejaran bastante que desear. Dado un número real $\varepsilon > 0$ *adecuado* a A, es decir, tal que las bolas abiertas en $\mathbb{C}$ centradas en los distintos valores propios de A y de radio ε fuesen disjuntas dos a dos, existía un número real $\delta > 0$, que dependía de ε tal que si $\|A' - A\| < \delta$, entonces el espectro de A' estaba contenido en la unión de las anteriores bolas de radio ε. Esta era una de las formas de parafrasear dicha continuidad en términos de la dialéctica aritmética de δ contra ε.

Mi tesis doctoral (1978) trató el problema "soviético" de la estructura de *matrices que conmutan con su derivada*. Me fue propuesto en 1975 por mi director de tesis, el profesor José Juan Rodríguez Cano. El abordaje de este problema me confrontó casi de continuo con la dificultad inherente a la inestabilidad de la forma canónica de Jordan de una matriz. Útil fue la representación de los valores propios de una función matricial $B(t)$ mediante una integral paramétrica de contorno en el campo complejo

$$\lambda_j(t) = \frac{1}{m_j} \operatorname{tr} \left[\frac{1}{2\pi \mathrm{i}} \oint_{\Gamma_j} z(zI - B(t))^{-1} \, \mathrm{d}z \right].$$

El problema de discernir alguna caracterización de las funciones matriciales $B(t)$ dependientes de un parámetro real t que variaba en un intervalo (a,b), donde las funciones $b_{ij}(t)$ eran derivables y tales que $B(t)B'(t) = B'(t)B(t)$, era francamente evanescente. Perseguí a un fantasma en la niebla. El principal trabajo previo que leí fue el de Bogdanov y Chebotarev de 1959. Proponían estudiar el problema bajo la hipótesis de que la función matricial $B(\cdot)$ fuese *conservativa*, i.e. de que conservara constante la estructura de Jordan a lo largo del intervalo. Luché infructuosamente con toda forma canónica visible en el horizonte sobre matrices que dependían de parámetros, la de Arnold incluida. Hasta llegué a explorar qué información arrojaban las fórmulas de Frenet-Serret sobre las curvaturas de $B(t)$ mirada como una curva en $\mathbb{R}^{2n^2}$, dado el carácter local de la condición $B(t)B'(t) = B'(t)B(t)$.

En 1981, animado por Ion Zaballa, entré en contacto con Graciano de Oliveira y el grupo de matemáticos portugueses de la Universidad de Coimbra y de la Universidad de Lisboa que trabajaban bajo su liderazgo. Fue una sorpresa conocer los *valores singulares de matrices*, así como su entrelazamiento, y la intrigante analogía que guardaban con los factores invariantes de matrices polinómicas. Llegó a tenerse claro que el *entrelazamiento* era una obsesión conimbricense. Sin darme cuenta, me había topado con él en mi tesis al considerar la conmutación de $B(t)$ con $B'(t)$ sobre un subespacio vectorial de $\mathbb{C}^n$. Muy importante fue la visita que Graciano de Oliveira realizó al Colegio Universitario de Álava en mayo de 1982, en cuyas conferencias planteó de forma meridianamente clara los problemas de determinar las posibles formas canónicas de Jordan de matrices parcialmente prescritas. Fue una verdadera epifanía joyceana.

Cuando hacia 1983 vi el artículo de Markus y Parilis sobre el cambio de la forma canónica de Jordan de una matriz A cuando ésta sufría pequeñas perturbaciones, pensé que eso era lo que llevaba tanto tiempo buscando. Y pasé unos años fascinado con sus resultados y profundizando en su estudio y análisis. Pudimos extender estos resultados a la forma canónica de Brunovsky para sistemas lineales con control. También los extendimos a haces rectangulares de matrices. ¿Qué valor tenían los posibles radios r de los entornos $B(A, r)$ cuya existencia garantizaban los teoremas de perturbación de Markus y Parilis? Así, pasé de la teoría de perturbación espectral cualitativa a la teoría cuantitativa. Pero la transición sería lenta. De siempre el cálculo numérico me había parecido un asunto muy *raro*. Como de costumbre, otros temas también reclamaban mi atención.

En todo momento la ecuación matricial de Sylvester $AX - XB = C$ atrajo mi interés. Su caso particular $AX - XA = O$ estaba conectado con la cuestión de caracterizar el conjunto $\mathcal{C}(A)$ de las matrices que conmutaban con A. La aplicación $A \mapsto \mathcal{C}(A)$ en un marco topológico adecuado era continua en A_0 si y sólo si la matriz A_0 era *no derogatoria*. Era conocida la forma canónica de Jordan de la aplicación lineal $\Phi_{AB}(X) := AX - XB$, y que su núcleo sólo contenía a la matriz nula si y sólo si los espectros $\Lambda(A)$ y $\Lambda(B)$, de A y B, eran disjuntos. En tal caso, ¿cuál era el mínimo valor posible de $\|E\| + \|F\|$ para que las perturbaciones E y F satisficieran que $\Lambda(A+E) \cap \Lambda(B+F) \neq \emptyset$? No por casualidad, Varah denotó mediante $\operatorname{sep}(A, B)$ a esta cantidad y la llamó *la separación* de A y B. No fue ninguna chiripa que $\operatorname{sep}(A, B)$ fuese mayor o igual que el último valor singular de la aplicación lineal Φ_{AB}. Llamando $\nu(A, B)$ a la dimensión de $\operatorname{Ker} \Phi_{AB}$, no menor impacto me produjo el conocimiento del criterio $\nu(A, A) = \nu(A, B) = \nu(B, B)$ para la semejanza de las matrices A y B. Si B estaba suficientemente cerca de A —¿cuánto?— la igualdad $\nu(A, A) = \nu(A, B)$ bastaba para garantizar la semejanza de A y B.

Pronto se nos hizo patente el interés del estudio de la perturbación de los subespacios invariantes de una matriz A al cambiar ligeramente esta matriz: la famosa caracterización de los subespacios A-invariantes *estables*. Otra vez la integral de contorno compleja

$$\frac{1}{2\pi \mathrm{i}} \oint_{\Gamma_0} (zI - A)^{-1} \, \mathrm{d}z,$$

a la que yo llamo *el invariante de I. Gohberg*, nos daba el proyector de Riesz sobre el subespacio radical, $\mathscr{R}_{\lambda_0}(A)$, de A asociado al valor propio λ_0, donde el contorno Γ_0 rodeaba a este valor propio y dejaba fuera a los restantes valores propios de A. Esta representación permitía deducir la continuidad de la aplicación $A \mapsto \mathscr{R}_{\lambda_0}(A)$, i.e. que el subespacio invariante $\mathscr{R}_{\lambda_0}(A)$ era estable. Los demás subespacios invariantes estables no quedaban muy lejos. Si se denotaba por $\operatorname{Inv}(A)$ el retículo formado por los subespacios invariantes de A, la aplicación $A \mapsto \operatorname{Inv}(A)$ era continua en A_0 si y sólo si A_0 era una matriz no derogatoria. Otra vez las matrices no derogatorias. Una cierta sensación de *déjà vu* me invadía al leer que si Ω era un dominio simplemente conexo de $\mathbb{C}$ y $A(z)$ era una función matricial analítica definida sobre él, entonces

Inv($A(z)$) dependía analíticamente de $z \in \Omega$ si y sólo si $A(z)$ tenía constante la estructura de Jordan; es decir, si $A(\cdot)$ era conservativa, que hubieran dicho Bogdanov y Chebotarev.

El hecho de no poder calcular exactamente la forma canónica de Jordan, excepto en ejercicios académicos, prefabricados, me producía —me produce— verdadera insatisfacción. Con el tiempo, hacia 1990, aparecieron trabajos que encaraban esta cuestión de frente. El profesor José Antonio Días da Silva, de Lisboa, me había dicho más de una vez: "Cuando no sepas qué hacer con los valores propios de una matriz mírale el digrafo". El digrafo, el grafo dirigido, era el objeto a considerar. A fin de cuentas, una matriz $A = (a_{ij})$ no era más que un *grafo dirigido ponderado*, donde el *peso* de la arista (i, j) era el número a_{ij}. Para fijar las ideas, supongamos que de una matriz cuadrada M conocemos solamente con certeza que en determinadas posiciones tiene ceros. En las restantes posiciones pondremos un interrogante. De salida, M será una *matriz* cuyos elementos pertenecen al conjunto $\{0, ?\}$, donde "?" significa que no sabemos qué número figura ahí. A partir de esta información ¿qué podemos decir sobre la forma canónica de Jordan de las matrices de este modelo M? Dicho en términos más precisos, ¿cuáles son los tamaños de los bloques de Jordan asociados al valor propio 0? Existe un respuesta a esta cuestión, dada en términos de un digrafo, $D(M)$, asociado al modelo M de matriz. Esta respuesta afirma que para *casi todas* las sustituciones de los interrogantes "?" de M por números complejos arbitrarios, las matrices de este modelo M tienen la misma característica de Segré asociada a 0 y los restantes valores propios son simples y no nulos. Obviamente, estos valores propios dependerán de los valores complejos asignados a los interrogantes "?" en M. La *forma canónica genérica de Jordan* de M vendrá determinada por el digrafo $D(M)$. Este resultado junto con los teoremas de perturbación de Markus y Parilis nos permiten deducir información, exactamente calculable, sobre la forma de Jordan de toda matriz concreta con modelo de ceros seguros prefijado. Seguros, pues en algunos interrogantes podría haber ceros también. Así pues, se puede sostener que para matrices parcialmente conocidas se puede hallar de forma exacta una parte de su forma de Jordan.

Tras algunos años con las anteriores obsesiones combinatorias, hacia 1992 empecé a estar interesado en el problema concreto de hallar la distancia, respecto de la norma espectral, de una matriz A cuyos valores propios fueran todos simples al conjunto de matrices $X \in \mathbb{C}^{n\times n}$ que tuvieran algún valor propio múltiple. Más tarde este problema acabaría llamándose de Wilkinson. Desde el análisis matemático elemental parecía un problema de extremos condicionados: Denotando por $p_X(\lambda)$ el polinomio característico de la matriz X, el problema de Wilkinson quedaba reducido al problema de minimizar $\|X - A\|$ sujeto a la condición $R(p_X, p_X') = 0$ en el espacio $\mathbb{R}^{2n^2} \equiv \mathbb{C}^{n\times n}$: Problema resuelto. Pero, para quienes habían conocido la presencia de los valores singulares en casi todo problema de buscar matrices más cercanas a una dada que tuvieran una determinada característica espectral, la anterior solución —caso de serlo— no era satisfactoria. Cuando apareció la solución de Malyshev (1999) a este problema, dada mediante optimización global del penúltimo valor singular de una matriz dependiente de un parámetro real, —una vez más— estaba preparado para reconocerla.

Así tropecé con el concepto de *pseudoespectros* de una matriz A. Dado un número real $\delta > 0$ ahora se trataba de ver hasta qué punto se alejaban del espectro $\Lambda(A)$ de A los espectros $\Lambda(X)$ de las matrices X situadas en el entorno $\|X - A\| < \delta$ de A. Se habían invertido los papeles de ε y δ de la definición de la continuidad del espectro $\Lambda(A)$ como función de A. Ahora primero fijábamos un $\delta > 0$. Si fuera suficientemente pequeño los valores propios de X no deberían estar muy lejos de los de A. Pero, esta cuestión inversa no era tan fácil, pudiendo estar los valores propios ξ de X muy alejados de los correspondientes valores propios α de A aun para valores de δ relativamente pequeños. Todo dependía de la sensibilidad (número de condición) del valor propio α a los cambios experimentados por A al transformarse en X. ¡Cómo ha cambiado nuestra forma de trabajar en matemáticas! Hace años que sin ayuda de los programas matemáticos de ordenador no sabría por dónde empezar. Quiero loar al programa MATLAB y a todos sus constructores. ¿Qué podríamos hacer —algunos— sin él? Aunque tengamos de origen una prevención contra el cálculo numérico. ¿Que qué me gusta de Matlab? Pues que calcula los valores singulares de matrices maravillosamente y permite conjeturar acertadamente los rangos de muchas de ellas. Quiero acabar dando las gracias más sentidas a Javier Asencor y a José María González de Durana, mis *maestros* de informática matemática. Auténticos magos de la conjetura simulada. El primero veía a π y sus parientes hasta con el más raro disfraz. El segundo, encontraba discontinuidades ocultas en superficies aparentemente lisas.

Por último lo más importante. El método de trabajo. Estudias tu solo. No hablas con nadie de tus dudas. Cuando has escrito la tesis, la presentas. Nada en el ambiente universitario en el que estudié parecía contradecir este proceder. No cabía pensar en trabajos colectivos. El trabajo matemático era —necesariamente— una tarea individual. Así llegamos hasta 1975. A partir de 1981 hubimos de hacer virtud de la necesidad. Era imperioso preguntar dudas a los compañeros. De este modo, profesores de asignaturas de matemáticas instrumentales para carreras anuméricas comenzamos a hacer seminarios semanales —que continúan al día de la fecha. En ellos pusimos en común dudas y problemas. Con el tiempo nos autodenominamos grupo de álgebra lineal; más tarde —cuando desde la nomenclatura pintaban bastos— mutamos en grupo de análisis matricial para no herir susceptibilidades. El nombre es lo de menos. Sin el auxilio de mis compañeros este camino a la perturbación hubiese abortado, más pronto que tarde. Con la ayuda de Graciano de Oliveira y demás matemáticos de Coimbra descubrimos que la realización de seminarios no era una herramienta más: era *la herramienta por antonomasia*. Miles de gracias a todos ellos. La disyuntiva ¿estudias o investigas? —que claramente me advirtiera Graciano de Oliveira allá por 1983— tardé mucho en asimilarla. Por mi educación e inclinaciones pertenezco a la *escuela contemplativa*. El ejemplo de mis compañeros de la *escuela activista* me ha ido llevando al método adecuado: el problema te habla, hay que saber escucharlo. Ya lo dijo el refrán: Por el hilo, se saca el ovillo.

Acknowledgements: Work (partially) supported by the Spanish Ministry of Education and Science Project MTM2010–19356–C02–01, and the Basque Government Projects GIC10/169–IT–361–10 and GIC13/IT–710–13.